MGMT methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status

Bady, P ; Sciuscio, D ; Diserens, A C ; Bloch, J ; van den Bent, M J ; Marosi, C ; Dietrich, P Y ; Weller, M ; Mariani, L ; Heppner, F L ; McDonald, D R ; Lacombe, D ; Stupp, R ; Delorenzi, M ; Hegi, M E

Abstract: The methylation status of the O(6)-methylguanine-DNA methyltransferase (MGMT) gene is an important predictive biomarker for benefit from alkylating agent therapy in glioblastoma. Recent studies in anaplastic glioma suggest a prognostic value for MGMT methylation. Investigation of pathogenetic and epigenetic features of this intriguingly distinct behavior requires accurate MGMT classification to assess high throughput molecular databases. Promoter methylation-mediated gene silencing is strongly dependent on the location of the methylated CpGs, complicating classification. Using the HumanMethylation450 (HM-450K) BeadChip interrogating 176 CpGs annotated for the MGMT gene, with 14 located in the promoter, two distinct regions in the CpG island of the promoter were identified with high importance for gene silencing and outcome prediction. A logistic regression model (MGMT-STP27) comprising probes cg1243587 and cg12981137 provided good classification properties and prognostic value (kappa = 0.85; log-rank $p < 0.001$) using a training-set of 63 glioblastomas from homogenously treated patients, for whom MGMT methylation was previously shown to be predictive for outcome based on classification by methylation-specific PCR. MGMT-STP27 was successfully validated in an independent cohort of chemoradiotherapy-treated glioblastoma patients (n = 50; kappa = 0.88; outcome, log-rank $p < 0.001$). Lower prevalence of MGMT methylation among CpG island methylator phenotype (CIMP) positive tumors was found in glioblastomas from The Cancer Genome Atlas than in low grade and anaplastic glioma cohorts, while in CIMP-negative gliomas MGMT was classified as methylated in approximately 50 % regardless of tumor grade. The proposed MGMT-STP27 prediction model allows mining of datasets derived on the HM-450K or HM-27K BeadChip to explore effects of distinct epigenetic context of MGMT methylation suspected to modulate treatment resistance in different tumor types.

DOI: https://doi.org/10.1007/s00401-012-1016-2

CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status. Acta Neuropathologica, 124(4):547-560.
DOI: https://doi.org/10.1007/s00401-012-1016-2

MGMT methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status

Pierre Bady[1,2,3], Davide Sciuscio[1], Annie-Claire Diserens[1], Jocelyne Bloch[1], Martin J van den Bent[4], Christine Marosi[5], Pierre-Yves Dietrich[6], Michael Weller[7,8], Luigi Mariani[9], Frank L Heppner[10], David R Mcdonald[11], Denis Lacombe[12], Roger Stupp[1], Mauro Delorenzi[2,3,13], Monika E Hegi[1,13] on behalf of the investigators of the European Organisation for Research and Treatment of Cancer and the National Cancer Institute of Canada Clinical Trials Group.

[1]Department of Clinical Neurosciences, Lausanne University Hospital, Lausanne, Switzerland;
[2]Bioinformatics Core Facility, Swiss Institute for Bioinformatics, Lausanne
[3]Département de formation et de recherche ; Lausanne University Hospital,
[4]Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands
[5]Medical University of Vienna, Vienna, Austria
[6]University Hospital Geneva, Geneva, Switzerland
[7]Department of Neurology, University of Tübingen, Tübingen; Germany
[8]Department of Neurology, University Hospital Zurich, Zurich, Switzerland
[9]Department of Neurosurgery, Inselspital Berne, Berne, Switzerland
[10]Department of Neuropathology, University Hospital Zurich, Zurich, Switzerland
[11]Neurology and Neuro-Oncology, London Regional Cancer Program, London Health Sciences Centre, University of Western Ontario, London, Ontario, Canada
[12]EORTC Headquaters, Brussels, Belgium
[13]National Center of Competence in Research Molecular Oncology, ISREC-SV-EPFL, Lausanne, Switzerland

Keywords: MGMT, DNA methylation, MSP, Infinium methylation platform, prediction model

Running title: MGMT methylation prediction model

Grant support: This work was supported by the Swiss National Science Foundation 3100A0_122557/1 (MEH, MD), the National Center of Competence in Research (NCCR) Molecular Oncology (MEH), and the Brain Tumor Funders Cooperative (MEH).

Corresponding Author: Monika E. Hegi, Laboratory of Brain Tumor Biology and Genetics, Department of Neurosurgery, Centre Hospitalier Universitaire Vaudois (CHUV BH19-110), 46 rue du Bugnon, Lausanne 1011, Switzerland
Phone: +41-21-314-2582, Fax: +41-21-314-2587, Email: Monika.Hegi@chuv.ch

Abstract:

The methylation status of the O^6-methylguanine-DNA methyltransferase (*MGMT*) gene is an important predictive biomarker for benefit from alkylating agent therapy in glioblastoma. Recent studies in anaplastic glioma suggest a prognostic value for *MGMT* methylation. Investigation of pathogenetic and epigenetic features of this intriguingly distinct behavior requires accurate *MGMT* classification to assess high throughput molecular databases. Promoter methylation-mediated gene silencing is strongly dependent on the location of the methylated CpGs, complicating classification. Using the HumanMethylation450 (HM-450K) BeadChip interrogating 176 CpGs annotated for the *MGMT* gene, with 14 located in the promoter, two distinct regions in the CpG island of the promoter were identified with high importance for gene silencing and outcome prediction. A logistic regression model (MGMT-STP27) comprising probes cg1243587 and cg12981137 provided good classification properties and prognostic value (kappa=0.85; logrank $p<0.001$) using a training-set of 63 glioblastoma from homogenously treated patients, for whom *MGMT* methylation was previously shown to be predictive for outcome based on classification by methylation-specific PCR. MGMT-STP27 was successfully validated in an independent cohort of chemo-radiotherapy treated glioblastoma patients (n=50; Kappa=0.88; outcome, logrank $p<0.001$). Lower prevalence of *MGMT* methylation among CpG island methylator phenotype (CIMP) positive tumors was found in glioblastoma from The Cancer Genome Atlas than in low grade and anaplastic glioma cohorts, while in CIMP-negative gliomas *MGMT* was classified methylated in approximately 50% regardless of tumor grade.

The proposed MGMT-STP27 prediction model allows mining of datasets derived on the HM-450K or HM-27K BeadChip to explore effects of distinct epigenetic context of *MGMT* methylation suspected to modulate treatment resistance in different tumor types.

Introduction

High throughput platforms for genome wide DNA methylation analysis have allowed establishing the methylome of large series of patient samples. Pattern analysis of respective datasets have identified CpG island methylator phenotypes (CIMP) for several tumor types such as colon cancer [19,21,42] and more recently also glioma (G-CIMP) [29,43,45]. However, classification of samples as being silenced by aberrant methylation for a given gene is not obvious, since the relationship between CpG-methylation at individual sites, the extent of the overall CpG island methylation, and their effect on gene silencing is strongly dependent on the location within the gene [46]. Promoter methylation of the repair gene O^6-methylguanine-DNA methyltransferase (*MGMT*) is a predictive factor for benefit from alkylating agent therapy in glioblastoma patients [3,16,33]. The predictive value of the *MGMT* status is supported by recent findings in two clinical trials, comparing radiotherapy versus temozolomide (TMZ) treatment. In these trials for elderly patients retrospective analysis of the *MGMT* methylation status was associated with prediction of good outcome in the TMZ-, but not the RT-arm [24,52]. Furthermore, the *MGMT* status has been prospectively validated in a phase III trial as biomarker for favourable outcome in glioblastoma patients treated with temozolomide [12]. Repair by MGMT reverses alkylation at the O^6-position of guanine, one of the most toxic lesions induced by alkylating agents such as temozolomide (TMZ), thereby blunting the treatment effect [18,30]. Hence, the *MGMT* methylation status has become a biomarker used for patient stratification or patient selection in clinical trials for glioblastoma patients [12,39,50]. Surprisingly, recent studies in anaplastic glioma (WHO grade III) suggest a prognostic value [44,51]. In order to investigate pathogenetic and epigenetic features associated with this intriguingly distinct behaviour of anaplastic glioma compared to glioblastoma, it is of high interest analyzing large datasets of glioma for which DNA methylome data has been reported, and classifying them by their *MGMT* gene promoter methylation status for integration into multi-dimensional molecular and clinical data analysis. Several glioma datasets comprising methylome data obtained on the Infinium HumanMethylation27 (HM-27K) or HM-450K BeadChip that interrogate at single-nucleotide resolution over 27,000 or 485,000 methylation sites per sample, respectively, have become publicly available [7,29,43,45]. The most comprehensive glioblastoma dataset with over 200 samples is from The Cancer Genome Atlas [29,40], widely used for hypothesis generation and validation in brain tumor research [17,48,53], however, the *MGMT* methylation status has not yet been annotated.

The objective of this study was to propose a model determining the probability of *MGMT* promoter methylation allowing classification into methylated and unmethylated samples based on CpG methylation data obtained on the widely used HM-450K or HM-27K BeadChip. The model can be applied to other datasets for example to further investigate of the relationship of *MGMT* methylation with CIMP and other molecular and clinical parameters. The basic idea was to train the model using methylation-specific PCR (MSP) based classification that we have shown to predict favorable outcome in a homogenously and prospectively treated glioblastoma patient population and for which we have obtained HM-450K data [15,16]. Standard treatment included the alkylating agent temozomoide (TMZ) concomitant with and adjuvant to radiotherapy [36,38]. At the same time this study allowed investigation of the relationship between location specific CpG methylation, *MGMT* gene expression and outcome, supporting the mechanistic hypothesis that methylation dependent gene silencing results in loss of expression and subsequently benefit from alkylating agent therapy in glioblastoma.

Material & methods

Patient samples and external data sets

DNA methylation profiles were established for 63 glioblastoma tissues from 59 patients and five non-tumoral brain tissues (epilepsy surgery). All glioblastoma patients were treated within a phase II or a randomized phase III trial [36,38] and provided written informed consent for molecular studies of their tumors. The protocols were approved by the ethics committees at each participating center and the respective competent authorities. For this patient cohort (M-GBM) that served as training set, detailed clinical information, treatment [37], and molecular data was available, including gene expression data [26], and the *MGMT* methylation status based on classic methylation-specific PCR (MSP) [15,16].

Four external glioma DNA methylation datasets associated with clinical information were used: The first, prospectively collected glioblastoma samples of a cohort of 50 patients (E-GBM) treated with combined chemo-radiotherapy with TMZ (Stupp protocol) for which HM-27K and methylation-specific pyrosequencing-based (MS-PSeq) *MGMT* methylation information was available [7] (see supplementary Figure S1 for location of 5 interrogated CpGs). The second dataset with HM-27K information consisted of 241 glioblastoma samples (TCGA-GBM, survival information available for 239 samples) and was downloaded from

The Cancer Genome Atlas (TCGA) website (http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) [29,40]. The third comprised HM-27K data from 67 anaplastic glioma (WHO grade III) (VB-Glioma-III) [45] of a cohort of homogenously treated patients, including *MGMT* methylation classification based on methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) (see Figure S1 for location of the 3 interrogated CpGs, probes used are described [44]). The fourth dataset comprised 71 low grade and anaplastic glioma samples (29 WHO grade II and 42 grade III) profiled on the HM-450K (T-Glioma-II/III) [43].

DNA methylation analysis

DNA was isolated from frozen tissues whereof 1.0 µg DNA was converted by bisulfite using the EZ DNA Methylation™ Kit (# D5001 Zymo Research Corporation) according to the manufacturer's instructions. DNA methylation analysis was performed on the HM-450K (Illumina) as recommended at the Genomics platform at the University of Geneva. For each interrogated CpG two site-specific probes are present, one designed for the methylated and another for the unmethylated locus to which the chemically converted DNA gets hybridized. Single-base extension of the hybridized probes incorporates a labeled ddNTP, which allows subsequent quantification of methylated and unmethylated alleles (http://www.illumina.com/technology/infinium_methylation_assay.ilmn).

Data analysis

The intensities of methylated and unmethylated signals were normalized using the Illumina GenomeStudio program. In the annotation file 176 CpG probes were associated with the *MGMT* gene, whereof 25 are shared with the HM-27K that was used in three of the four external datasets. The DNA methylation information was summarized by M-values as recommended by Du et al. [4]:

$$M-values = log2\left(\frac{\max(signalB,0)+1}{\max(signalA,0)+1}\right)$$

The terms 'signal A' and 'signal B' correspond to the intensities of the unmethylated and methylated probes.

Statistical methods

The relationship between the methylation status of the *MGMT* promoter defined by MSP and the probes located in the *MGMT* promoter region (CHR10: 131264700-131266300, genome build 37) (Figure 1, S1) present on the HK-450K was evaluated by logistic

regressions [25,28]. A two-step procedure was used to construct the model to calculate the probability of *MGMT* promoter methylation for subsequent binary classification. In the first step, all univariate models were tested and only probes significantly associated with the MSP classification were selected using the loglikelihood ratio test (LRT) and Bonferroni correction for multiple testing. In a second step, we determined the optimal model built by stepwise logistic model building based on the corrected Akaike's criterion (AICc) [1,20,47] that limits overfitting. Two models were constructed, one with all selected probes, and one using only the probes common to both, the HM-27K and the HM-450K platforms. The model performance was assessed by internal-validation based on a bootstrap procedure with optimism/bias correction [13]. This procedure validates the process used to fit the original model and it provides a bias value defined by the difference between the index from the original dataset and the average of indices from the resampling procedure (200 repetitions). The Kappa index and the proportion of correct classification were used to evaluate the concordance between the observed and the predicted methylation status. The M-value distributions of the probes selected by the model were compared by pairwise quantile-quantile representation (QQ-plot) test for the five datasets and non-parametric Smirnov-Kolmogorov test for equal distribution.

For the external datasets, the probability of *MGMT* promoter methylation and subsequent classification were determined using the model compatible with the HM-27K platform. The probability values were associated with Wald-based confidence intervals (CI 95%) to evaluate the uncertainty of the model [9]. CIMP positive tumors were identified using unsupervised clustering methods similar to Noushmehr et al. and Turcan et al. [29,43]. The relationship between the predicted *MGMT* status and presence of CIMP was assessed by chi-squared tests with p-values computed by Monte Carlo simulation, because cell counts were expected to be inferior to five [31].

The Kaplan-Meier (KM) curves and log-rank tests were estimated for each dataset and predictor [13,41]. Univariate and multivariate survival models were assessed using the Cox proportional hazards regression model and the log-likelihood ratio test (LRT).

Analyses and graphical representations were performed using R-2.14.2 and the R packages design and survival [32].

Results

Calibration model to predict the methylation status of the *MGMT* promoter using the HM-450K platform

The prediction model is based on a glioblastoma patient cohort treated with standard chemo-radiotherapy within two prospective clinical trials. The *MGMT* promoter methylation status of this cohort was available from classic MSP discriminating methylated and unmethylated samples. Most importantly, the results of this MSP-test have been shown to be of predictive value for benefit from the addition of temozolomide chemotherapy [15,16,26,37]. The results of the MSP assay were considered as reference for model construction. A total of 63 glioblastoma samples (M-GBM) and five non-tumoral brain tissues were analyzed on the HM-450K beadarray platform. The five non-tumoral brain tissues were classified as *MGMT* unmethylated based on MSP. The clinical and molecular information of patient samples is summarized in Table S1. Of the 176 CpG probes on the HM-450K annotated as mapping to the *MGMT* gene, 14 are in the CpG island located in the *MGMT* promoter region encompassing the transcription start site (TSS) (Figure 1). This CpG island has been shown to be determinant for regulation of expression [8,23,27,35]. The remaining CpGs are mostly found in the *MGMT* gene body (Supplementary Figure S2). The detection call rate for these 14 CpG island-associated probes was equal to 1.00, indicating reliable detection.

Association of probes with MSP, *MGMT* expression, and outcome

Eleven probes located in the TSS-encompassing CpG island of the *MGMT* promoter were significantly associated with the previously MSP-defined *MGMT* methylation status (Figure 1). We observed two prediction peaks. The strongest association was reached by the probes **cg12434587** and **cg12981137,** respectively (No. 10 and 16 in Figure 1). Of note, the CpG tested by the probe **cg12981137** (CHR 10: 131265575, genome build 37) is also interrogated by the reverse primer of the MSP assay (Supplementary Fig. S1) [6]. Strikingly, the highest negative correlation between methylation and gene expression (estimated by Affymetrix U133plus2, [26]) was observed for the same two probes (Spearman correlation coefficients -0.543 and -0.571, respectively). The negative correlation is consistent with promoter methylation mediated down-regulation of *MGMT* expression. Similarly, these two probes were the most negatively correlated with *MGMT* gene expression in the E-GBM and

the TCGA-GBM dataset for which expression data was available (Supplementary Fig. S3). The association of *MGMT* expression and DNA methylation at individual CpGs is visualized in a scatter plot for the M-GBM data set including non tumoral brain tissues in the Supplementary Figure S4. Remarkably, the probes, in vicinity to the TSS (probes 12-14 in Figure 1) were at best weakly associated with the MSP results (Figure 1). The strongest association with overall survival (OS) peaked at the same two probes, **cg12434587** and **cg12981137,** with p-values of $<10^{-5}$ and $<10^{-4}$, while again no association was found at CpGs in the vicinity of the TSS, and the 5'and 3' edges of the CpG island (Figure 1). The p-values for OS prediction of these two probes were similar to the one for the MSP-based methylation classification ($p<10^{-5}$). The association of OS with methylation at individual CpGs is shown for all datasets in Figure S3. Taken together, the CpG methylation probes (8, 9, 10 and 16) most correlated with expression also correspond to the probes highly associated with survival, and most correlated with MSP-based *MGMT* methylation prediction.

In order to test if other methylated CpGs are relevant for *MGMT* silencing we investigated the remaining CpGs, mainly located in the gene body. No negative correlation was observed between methylation and expression and no association with outcome (Supplementary Figure S2).

Stepwise model building

We aimed at building an optimal methylation predictor using multiple probes. In the first step, the 11 probes significantly associated with MSP-defined methylation were selected. The prediction performances with individual probes were very high. Expectedly, the highest sensitivity and specificity values were obtained for the probes cg12981137 and cg12434587 (0.906 and 0.944; 0.875 and 0.944) (Figure 1). Nonetheless, stepwise model building (see methods) indicated that MSP assay classification could be predicted even better using two or more probes.

The best model based on probes shared with the HM-27K platform (see Figure 1) comprised the two probes cg12434587 and cg12981137 (MGMT-STP27 model). The adjustment of the model was of high quality (supplementary Table S2). The sensitivity and specificity were equal to 0.969 and 0.889 respectively (Figure 2). The proportion of correct classification (0.926), the Kappa index (0.853), and the AUC (0.974) confirm the good performance of the model. The hazard risks (HR) based on MSP and the MGMT-STP27 model were similar (logrank test for both, $P<0.001$; MSP, HR = 0.229, 95% confidence

interval (CI 95%) = [0.115, 0.454]; MGMT-STP27, HR=0.277, CI 95% = [0.145, 0.529]) (Figure 2). The equation for this model is given below:

$$logit(y) = 4.3215 + 0.5271 \cdot cg12434587 + 0.9265 \cdot cg12981137$$

The methylation probability y can be computed by using the inverse logit function. For classification, we used a probability cut-off of 0.358, which empirically maximized the sum of sensitivity and specificity (supplementary Table S2; annotation of M-GBM sample classification supplementary Table S1). The classification by STP27 and MSP is visualized in Figure 2a. Furthermore, the classification by STP27 is projected onto the scatter plots comparing expression and DNA methylation at individual CpGs, reclassification from MSP-based prediction are indicated (Supplemental Figure S4).

The second model, obtained by using all 11 probes available on the HM-450K platform contained 4 probes: the two probes also selected in MGMT-STP27 (cg12981137, cg12434587) plus cg02022136 and cg23998405. The improvement over MGMT-STP27 is only marginal (inferior to 1 unit of the AICc, supplementary Table S2). However, the latter model showed a high inflated variance factor (8.9, supplementary Table S2) that reflects a problem of multi-colinearity. Consequently, this model was not considered for further analyses.

The internal validation based on the bootstrap procedure showed that the model MGMT-STP27 was relatively stable. The unbiased diagnostic accuracy (proportion of correct classification) was estimated to 91.24%, against 92.65% initially computed on the original dataset. For the Kappa index, the difference between unbiased and original value was only equal to 0.03 units (supplementary Table S3).

External validation of MGMT-STP27

We validated the use of the MGMT-STP27 model in an external data-set of 50 glioblastoma (E-GBM) analyzed on the HM-27K [7]. We dichotomized the MS-PSeq information of the *MGMT* promoter available for 47 cases into methylated and unmethylated. The cut-off was estimated at 7.28% average methylation based on a fitted regression model visualized in supplementary Figure S5 which is similar to previous reports applying a cut-off of 8% using the same MS-PSeq kit [10,33]. The correspondence between the predicted status of *MGMT* promoter methylation using MGMT-STP27 and MS-PSeq information was very high, as visualized in Figure 3a with a proportion of correct classification of 0.936, a Kappa index of 0.875, and sensitivity and specificity equal to 0.931 and 0.944, respectively. Outcome

prediction using MS-PSeq information or MGMT-STP27 was as follows: p=0.019, HR = 0.454, CI 95% [0.232; 0.891] for MS-PSeq, and p< 0.001, HR = 0.305, CI 95%= [0. 156; 0.596] for MGMT-STP27) (Fig. 3).

Next we tested the model in a dataset from a cohort of anaplastic oligodendroglioma and anaplastic oligoastrocytoma (WHO grade III) (VB-Glioma-III, n=67) generated with the 27K-platform [45]. The methylation status of the *MGMT* promoter was available from MS-MLPA for most cases (n=62). We observed a good concordance (Fig. 3d) with a good classification rate of 0.839, a Kappa index of 0.628, and sensitivity and specificity equal to 0.975and 0.609, respectively. As reported by van den Bent et al. [45], *MGMT* methylation as determined by MS-MLPA was significantly associated with favorable outcome of the patients with a p-value of 0.031 (log-rank test) (HR= 0.527, CI 95%=[0.292, 0.952]). This is similar to the prediction by our MGMT-STP27 model with a p-value of 0.029 (HR= 0.500, CI 95%=[0.266, 0.941]) as visualized in Figure 3.

Prediction of *MGMT* methylation status in TCGA-GBM cohort and a glioma grade II & III data set

The good performance of the MGMT-STP27 model in the two external data-sets suggests that we can appropriately predict the *MGMT* methylation status using common probes between the HM-450K and HM-27K platform. Prediction of the *MGMT* methylation status in 241 glioblastoma available from TCGA (TCGA-GBM) with HM-27K data, revealed a methylation frequency of 50% (120/241) (see annotation of samples in supplementary Table S4 that also includes more recent samples analyzed on the HM-450K platform) similar to our M-GBM cohort. Patients from the TCGA-GBM dataset (n=239) with a *MGMT* methylated glioblastoma had a more favorable outcome (p-value=0.047, log-rank test, Figure 4). The HR for the predicted methylation status was equal to 0.737 (CI 95% = [0.545, 0.997]).

The prediction of the *MGMT* methylation status using MGMT-STP27 in a cohort of grade II and III gliomas (T-Glioma-II/III) with HM-450K data, determined a methylation frequency of 79% (32/42) in grade III glioma and 86% (25/29) in grade II. The favorable outcome associated with *MGMT* methylation was confirmed as visualized in Figure 4.

Associations of *MGMT* methylation across tumor grades of glioma

Next we asked whether *MGMT* is part of the genes associated with CIMP using the TCGA-GBM data set [29]. Glioblastoma with a methylated *MGMT* promoter were significantly enriched among CIMP cases (15/20, 75% against 105/221, 48%; p-value = 0.023, Table S5).

However, the two *MGMT* probes selected in our prediction model do not cluster with the CIMP core genes as visualized in the heatmap of supplementary Figure S6. Classification of the glioblastoma into the three methylation clusters (CIMP+, and 2 non-CIMP clusters) as published by Noushmehr et al. [29] were available for 81 samples with HM-27K DNA methylation (Fig. 5, Supplementary Fig. S6). The 10 CIMP+ samples defined by the authors were all in the CIMP cluster obtained by unsupervised clustering methods in the present study. The distribution of samples with methylated *MGMT* was not significantly different for the two non-CIMP methylation clusters (p=0.214, Monte Carlo simulation) defined by Noushmehr et al. [29].

No association was found between the *MGMT* status and any of the expression-based glioblastoma subtypes proposed by Verhaak et al. [48] (p= 0.518, Monte Carlo simulation) (Fig. 5; Table S5).

In contrast, in both the VB-Glioma-III and the T-Glioma-II/III datasets basically all tumors clustering together displaying CIMP were classified as *MGMT* methylated by the MGMT-STP27 model (32/32 and 48/49), while the CIMP negative tumors exhibited a methylation frequency of 54% (19/35) and 50% (11/22) that is similar to the three glioblastoma datasets (Figures 5 & 6). Hence, the association of methylated *MGMT* with CIMP was much stronger in low grade and anaplastic glioma than in the TCGA-GBM. The difference is highly statistically significant (chi-squared test with p-value estimated by Monte Carlo simulation; 32/32 vs 15/20, p-value < 0.005 for VB-Glioma-III and 48/49 vs 15/20, p-value < 0.005 for T-Glioma-II/III). In the VB-Glioma-III dataset this association was retained when using the CIMP classification published by van den Bent and colleagues (*MGMT* methylated among CIMP+, 30/30; Figure 5).

Comparisons of M-value distribution across datasets and platforms

The pairwise comparisons of the M-value distribution for the two probes cg12434587 and cg12981137 across datasets revealed that M-values were generally higher in the grade II/III glioma (VB-Glioma-III, T-Glioma-II/III) as compared to glioblastoma (M-GBM and TCGA-GBM) (p-values <0.05 from Kolmogorov-Smirnov tests, supplementary Figure S7). However, the comparisons between the glioblastoma datasets, analyzed on the HM-450K (M-GBM) and the HM-27K (TCGA-GBM), respectively, showed similar M-values distributions (p=0.260 and p=0.145, supplementary Figure S7). Likewise, the M-value distributions among the datasets of grade II and III gliomas analyzed on the HM-27K (VB-Glioma-III) and the HM-450K (T-Glioma-II/III), respectively, were similar (p=0.435 and p=0.233, supplementary

Figure S7). Hence, indicating that differences observed between the studies were not due to a platform effect, but rather result from biological differences. Consequently, this could affect (bias) the prediction quality of the methylation status of *MGMT* for grade II and III glioma by MGMT-STP27, and may explain the higher number of positive calls by MGMT-STP27 when compared to MS-MLPA (specificity of only 0.609 when using MGMT-STP27 as predictor of MS-MLPA methylation calls). However, outcome prediction was equally good for both methylation call methods in the VB-Glioma-III data (Fig. 3).

Discussion

The analysis of the *MGMT* gene in glioblastoma using HM-450K methylation data has shown a strong CpG location dependent effect on patient outcome. To our knowledge, this is the first report describing the spatial relationship of CpG methylation in the *MGMT* promoter and the gene body and outcome of patients treated with alkylating agent therapy. Two regions of methylated CpGs with strong association to patient survival were identified ($p < 0.0001$ after Bonferroni's correction) separated by a prediction minimum at the TSS. The two identified regions were also associated with the strongest negative correlation to *MGMT* gene expression, consistent with CpG methylation mediated gene silencing and consequent sensitization to alkylating agent therapy due to lack of MGMT mediated repair in this homogenously treated patient population. The two regions identified encompass the differentially methylated regions 1 and 2 (DMR1 and 2, Figure 1) proposed by Malley et al. [23] to be most relevant for gene silencing when methylated in glioblastoma cell lines and xenografts. Shah et al. [35] defined 3 relevant regions, of which R2 and R3 encompass DMR1 and 2, and methylation of two of these three regions were associated with favorable progression free survival in their population of 44 glioblastoma patients treated with RT and concomitant and adjuvant TMZ. Most importantly, the region interrogated for diagnostic purposes using MSP [5,16] overlaps with the CpGs associated with best outcome prediction identified here. In contrast, none of the 161 CpGs interrogated outside the CpG island, located mostly in the gene body, showed an association with outcome, or a negative correlation with *MGMT* expression (Supplementary Figure S2). However, we cannot exclude that other CpGs may be more relevant as the once described here, since the BeadChip array does not interrogate all CpGs of the CpG island encompassing the *MGMT* promoter (Supplementary Fig. S1).

The CpG methylation probes present on the HM-450K and HM-27K BeadChip identified to be most relevant for gene silencing and outcome allowed construction of a model for

prediction of the *MGMT* methylation status. The use of logistic regression provides a simple model to calculate the methylation probability for a new sample based on two probes. Its ability to compute confidence and prediction intervals [2,20] may be of particular interest for treatment decisions for patients whose tumors display methylation probabilities close to the cut-off (Figures 2-4). This allows application of a "safety margin" as we do in EORTC26082 (NCT01019434) that selects unmethylated glioblastoma patients only (cut-off set at lower bound of 95% CI using a quantitative MSP assay [49]), since it omits TMZ, thereby limiting the risk to withhold TMZ in patients who may potentially profit from it. The model's good performance is reflected in similar or improved prediction of OS as compared to the MSP-based classification or MS-PSeq-based prediction in the E-GBM validation set, in accordance with high values for good classification, Kappa-value, and sensitivity and specificity measures (Figures 2 & 3).

Our model can be used for both the HM-450K and the HM-27K BeadChip. The platform effect was very weak. A higher amplitude of the methylation signal was detected in the low grade and anaplastic glioma samples that may simply reflect the fact that in non-glioblastoma usually both *MGMT* copies are present, while in glioblastoma only one is methylated and the other one is lost due to the characteristic high frequency of deletions of chromosome 10 that reached 90% in the M-GBM samples [22]. Consequently, presence of two methylated *MGMT* copies will lead to an increase of the ratio methylated to unmethylated alleles. For the model, however, this may generate a bias in the estimation of *MGMT* methylation probabilities based on the MGMT-STP27 model for non-glioblastoma tumors. The estimation could be improved by determining new optimal parameters in this population. Nevertheless, despite these limitations classification using the MGMT-STP27-based outcome prediction of the VB-Glioma-III dataset was similar to the one reported by the authors who used another method of *MGMT* testing (Figure 3). Prediction of the MGMT status in the TCGA-GBM confirmed a favorable outcome for patients with *MGMT* methylation, although the effect was weaker than in our homogenously treated cohort (M-GBM) and the E-GBM cohort in which all patients were treated with combined chemo-radiotherapy comprising TMZ. This is not surprising, since most patients in the TCGA cohort had not (yet) been treated according to the current standard of care of combined chemo-radiotherapy (collection before 2005), and many different types of therapy were reported for the patients in the respective annotation file [29].

The annotation of the *MGMT* status in the TCGA-GBM dataset according to MGMT-STP27 allowed determining that *MGMT* is not a CIMP gene in glioblastoma although CIMP tumors were more likely to be *MGMT* methylated. Further, the prevalence of *MGMT*

methylated glioblastoma is not different in the 2 non-CIMP methylation clusters defined by Noushmehr et al. [29], nor are they enriched in any of the expression-based glioblastoma subtypes, suggesting that *MGMT* methylation is not associated with a particular pathogenetic mechanism involved in the development of de novo glioblastoma.

This is in contrast to grade II and III glioma (VB-Glioma-III and T-GliomaII/III) in which *MGMT* is methylated in basically all CIMP tumors according to our classification model. It has been proposed that *MGMT* methylation may represent an epiphenomenon of CIMP in the context of grade III glioma [45]. This association of with CIMP with *MGMT* methylation may provide the key to understand why *MGMT* methylation is associated with a prognostic and not a predictive value for benefit from alkylating agent containing chemotherapy in anaplastic glioma as suggested by two independent clinical trials [44,51]. Most of the VB-Glioma-III samples analyzed here in fact originate from one of these two studies and were characterized for CIMP [45]. Anaplastic gliomas with CIMP accumulate other known favourable prognostic factors such as mutations of the isocitrate dehydroxygenase (*IDH*) genes, 1p/19 co-deletions, and also *MGMT* promoter methylation, plus a plethora of other methylated genes whose contribution to response to therapy remains to be explored and exploited. It has become clear that these CIMP positive tumors represent a pathogenetically different disease driven by epigenetic alterations mediated in most cases by IDH1/2 mutations [11,43]. Interestingly, non-CIMP anaplastic gliomas showed a *MGMT* promoter methylation frequency similar to glioblastoma. It remains to be seen if in this CIMP negative patient subpopulation the *MGMT* status is predictive for benefit from alkylating agent therapy like in glioblastoma or has a prognostic value. This question can be addressed in the ongoing CATNON trial (EORTC 26053-22054; NCT00626990) for anaplastic glioma comparing radiation therapy with or without temozolomide. The same question applies to low grade glioma where radiation versus temozolomide treatment is tested (EORTC 22033-26033, NCT00182819) and the role of CIMP and *MGMT* methylation will need to be dissected. Since the HM-450K BeadChip allows the use of paraffin embedded tumors, comprehensive DNA methylation analysis of samples collected within these clinical trials has become feasible.

The proposed MGMT-STP27 *MGMT* classification model will allow investigation of distinct epigenetic features associated with *MGMT* silencing in the context of CIMP positive or CIMP negative gliomas by multidimensional analysis of respective molecular and clinical data. Such alterations likely modulate response to therapy and may be exploited for improvement of personalized therapy.

Conflict of Interest

MEH is an advisor to MDxHealth. MEH, MW, MJvdB, DRM, and RS have an advisory role and have received honoraria from MSD.

Acknowledgement

We thank all patients who participated in the study and provided informed consent for translational research on their tumor tissues. We acknowledge the great contributions of local pathologists, the physicians and nurses taking care of the patients. We thank Marie-France Hamou for excellent technical support and the genomics platform in Geneva directed by Dr Patrick Descombes for methylation profiling.

Translational research in this study was supported by the Swiss National Science Foundation 3100A0_122557/1 (MEH, MD), the National Center for Competence in Research (NCCR) Molecular Oncology (MEH, MD), and the Brain Tumor Funders Cooperative (MEH).

References

1 Burnham KP, Anderson DR (2002) Model selection and multimodel inference: A practical information theoretic approach. Springer-Verlag, New York

2 Collett D (2002) Modelling binary data. Chapman & Hall/CRC, London

3 Criniere E, Kaloshi G, Laigle-Donadey F et al (2007) MGMT prognostic impact on glioblastoma is dependent on therapeutic modalities. J Neurooncol 83:173-179

4 Du P, Zhang X, Huang CC et al (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 11:587

5 Esteller M, Garcia-Foncillas J, Andion E et al (2000) Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med 343:1350-1354

6 Esteller M, Hamilton SR, Burger PC et al (1999) Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res 59:793-797

7 Etcheverry A, Aubry M, de Tayrac M et al (2010) DNA methylation in glioblastoma: impact on gene expression and clinical outcome. BMC Genomics 11:701

8 Everhard S, Tost J, El Abdalaoui H et al (2009) Identification of regions correlating MGMT promoter methylation and gene expression in glioblastomas. Neuro Oncol 11: 348-56

9 Faraway JJ (2006) Extending linear models with R: Generalized linear, mixed effects and nonparametric regression models. Chapman & Hall/CRC, Boca Raton, FL

10 Felsberg J, Thon N, Eigenbrod S et al (2011) Promoter methylation and expression of MGMT and the DNA mismatch repair genes MLH1, MSH2, MSH6, and PMS2 in paired primary and recurrent glioblastomas. Int J Cancer 129:659-670

11 Figueroa ME, Abdel-Wahab O, Lu C et al (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18:553-567

12 Gilbert MR, Wang M, Aldape KD et al (2011) RTOG 0525: A randomized phase III trial comparing standard adjuvant temozolomide (TMZ) with a dose-dense (dd) schedule in newly diagnosed glioblastoma (GBM). J Clin Oncol 29:suppl; abstr 2006

13 Harrell FE (2001) Regression modeling strategies, with applications to linear models, survival analysis and logistic regression. Springer, New York

14 Harris LC, Potter PM, Tano K et al (1991) Characterization of the promoter region of the human O6-methylguanine-DNA methyltransferase gene. Nucleic Acids Res 19:6163-6167

15 Hegi ME, Diserens AC, Godard S et al (2004) Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. Clin Cancer Res 10:1871-1874

16 Hegi ME, Diserens AC, Gorlia T et al (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352:997-1003

17 Hegi ME, Janzer RC, Lambiv WL et al (2012) Presence of an oligodendroglioma-like component in newly diagnosed glioblastoma identifies a pathogenetically heterogeneous subgroup and lacks prognostic value: central pathology review of the EORTC_26981/NCIC_CE.3 trial. Acta Neuropathol

18 Hegi ME, Liu L, Herman JG et al (2008) Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in

glioblastoma and clinical strategies to modulate MGMT activity. J Clin Oncol 26:4189-4199

19 Hinoue T, Weisenberger DJ, Lange CP et al (2012) Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res 22:271-282

20 Hosmer DW, Lemeshow S (2000) Applied logistic regression. Wiley Interscience, Chichester, NY

21 Issa JP (2004) CpG island methylator phenotype in cancer. Nat Rev Cancer 4:988-993.

22 Lambiv WL, Vassallo I, Delorenzi M et al (2011) The Wnt inhibitory factor 1 (WIF1) is targeted in glioblastoma and has a tumor suppressing function potentially by induction of senescence. Neuro Oncol 13:736-747

23 Malley DS, Hamoudi RA, Kocialkowski S et al (2011) A distinct region of the MGMT CpG island critical for transcriptional regulation is preferentially methylated in glioblastoma cells and xenografts. Acta Neuropathol

24 Malmström A, Grønberg BH, Marosi C et al ((in press)) Temozolomide versus standard 6-week radiotherapy versus hypofractionated radiotherapy for patients aged over 60 years with glioblastoma: the Nordic randomized phase 3 trial. Lancet Oncol

25 McCullagh P, Nelder JA (1989) Generalized linear models. Chapman & Hall/CRC, London

26 Murat A, Migliavacca E, Gorlia T et al (2008) Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. J Clin Oncol 26:3015-3024

27 Nakagawachi T, Soejima H, Urano T et al (2003) Silencing effect of CpG island hypermethylation and histone modifications on O6-methylguanine-DNA methyltransferase (MGMT) gene expression in human cancer. Oncogene 22:8835-8844

28 Nelder JA, Wedderburn RWM (1972) Generalized linear models. Journal of the Royal Statistical Society, Series A 135:370-384

29 Noushmehr H, Weisenberger DJ, Diefes K et al (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 17:419-420

30 Ochs K, Kaina B (2000) Apoptosis induced by DNA damage O6-methylguanine is Bcl-2 and caspase-9/3 regulated and Fas/caspase-8 independent. Cancer Res 60:5815-5824

31 Patefield WM (1981) Algorithm AS159. An efficient method of generating r x c tables with given row and column totals. Applied Statistics 30:91-97

32 R Development Core Team (2011) R: A language and environment for statistical computing. Vienna, Austria

33 Reifenberger G, Hentschel B, Felsberg J et al (2011) Predictive impact of MGMT promoter methylation in glioblastoma of the elderly. Int J Cancer

34 Sciuscio D, Diserens AC, van Dommelen K et al (2011) Extent and patterns of MGMT promoter methylation in glioblastoma- and respective glioblastoma-derived spheres. Clin Cancer Res 17:255-266

35 Shah N, Lin B, Sibenaller Z et al (2011) Comprehensive analysis of MGMT promoter methylation: correlation with MGMT expression and clinical response in GBM. PLoS One 6:e16146

36 Stupp R, Dietrich P, Ostermann Kraljevic S et al (2002) Promising survival for patients with newly diagnosed glioblastoma multiforme treated with concomitant

radiation plus temozolomide followed by adjuvant temozolomide. J Clin Oncol 20:1375-1382

37 Stupp R, Hegi ME, Mason WP et al (2009) Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol 10:459-466

38 Stupp R, Mason WP, van den Bent MJ et al (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352:987-996

39 Stupp R, Van Den Bent MJ, Erridge SC et al (2010) Cilengitide in newly diagnosed glioblastoma with *MGMT* promoter methylation: Protocol of a multicenter, randomized, open-label, controlled phase III trial (CENTRIC). J Clin Oncol 28:15s, (suppl; abstr TPS152)

40 The Cancer Genome Atlas Consortium (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455:1061-1068

41 Therneau TM, Grambsch PM (2000) Modeling survival data: Extending the cox model. Springer, New York

42 Toyota M, Ahuja N, Ohe-Toyota M et al (1999) CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 96:8681-8686

43 Turcan S, Rohle D, Goenka A et al (2012) IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483:479-483

44 van den Bent MJ, Dubbink HJ, Sanson M et al (2009) MGMT promoter methylation is prognostic but not predictive for outcome to adjuvant PCV chemotherapy in anaplastic oligodendroglial tumors: A report from EORTC Brain Tumor Group Study 26951. J Clin Oncol 9:5881-5886

45 van den Bent MJ, Gravendeel LA, Gorlia T et al (2011) A hypermethylated phenotype in anaplastic oligodendroglial brain tumors is a better predictor of survival than MGMT methylation in anaplastic oligodendroglioma: a report from EORTC study 26951. Clin Cancer Res 17:7148-7155

46 van Vlodrop IJ, Niessen HE, Derks S et al (2011) Analysis of promoter CpG island hypermethylation in cancer: location, location, location! Clin Cancer Res 17:4225-4231

47 Venables WN, Ripley BD (2002) Modern applied statistics with S Springer, New York

48 Verhaak RG, Hoadley KA, Purdom E et al (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17:98-110

49 Vlassenbroeck I, Califice S, Diserens AC et al (2008) Validation of real-time methylation-specific PCR to determine O6-methylguanine-DNA methyltransferase gene promoter methylation in glioma. J Mol Diagn 10:332-337

50 Weller M, Stupp R, Reifenberger G et al (2010) MGMT promoter methylation in malignant gliomas: ready for personalized medicine? Nat Rev Neurol 6:39-51

51 Wick W, Hartmann C, Engel C et al (2009) NOA-04 randomized phase III trial of sequential radiochemotherapy of anaplastic glioma with procarbazine, lomustine, and vincristine or temozolomide. J Clin Oncol 27:5874-5880

52 Wick W, Platten M, Meisner C et al (2012) Temozolomide chemotherapy alone versus radiotherapy alone for malignant astrocytoma in the elderly: the NOA-08 randomised, phase 3 trial. The Lancet Oncology

53 Zinn PO, Majadan B, Sathyan P et al (2011) Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS One 6:e25451

Figure Legends

Figure 1. CpG methylation of the *MGMT* promoter region, *MGMT* expression and patient survival in M-GBM. (**a**) The Spearman and Pearson correlation between gene expression (probe 204880_at from Affymetrix U133plus2) and M-values of the 18 CpG methylation probes from the Infinium humanmethyltion 450K BeadChip (HM-450K) of the M-GBM cohort are visualized on a scale representing the physical location in the CpG island of the promoter region encompassing the transcription start site (TSS) (genome build 37). (**b**) The association between overall survival (OS) and CpG methylation of distinct probes are displayed (p-values, univariate Cox regression model and log-likelihood ratio test; p-values, minus-log10-transformed). The p-value for classification by MSP is indicated at its physical location (primer set, red). (**c**) The association between *MGMT* promoter methylation classification based on MSP and the 18 selected CpG methylation probes from the HM-450K are shown (logistic regression and log-likelihood ratio test; p-values, minus-log10-transformed). The dotted grey lines in **b** & **c** correspond to the threshold of 0.05. The graph at the bottom indicates the physical location of the TSS (TSS1, according to Harris et al. [14]; TSS2 according to gene build 19); the location of the CpG island/ individual CpGs, green; the differentially methylated regions 1 and 2, DMR1 & 2, as defined by Malley at al. [23] blue; the primers for MSP [6], red; the region for MS-clone sequencing in glioblastoma, MS-CSeq [34]; the CpGs interrogated by methylation specific multiplex ligation-dependent probe amplification, MS-MPLA, purple [44], and methylation-specific pyrosequencing, MS-PSeq, pink [7]. The names of the CpG probes interrogated on HM-450K are given on the right. Probes present on both platforms (HM-450K and HM-27K) are indicated by triangles, probes only present on the HM-450K are represented by squares. See supplementary Figure S1 for exact locations of CpGs interrogated by the different assays. The symbols are explained on the right hand side. We note that the CpG methylation probes (8, 9, 10 and 16) most correlated with expression also correspond to the probes highly associated with survival, and most correlated with MSP-based *MGMT* methylation prediction.

Figure 2. Performance of the stepwise logistic regression model (MGMT-STP27) for prediction of methylation status of the *MGMT* promoter. (**a**) displays the estimated probability of methylation against the logit-transformed response fitted for the M-GBM dataset. The observed values are given by full black squares, indicating same or different classification by STP27. Fitted values and their confidence intervals [CI] at 95%, estimated by simulation,

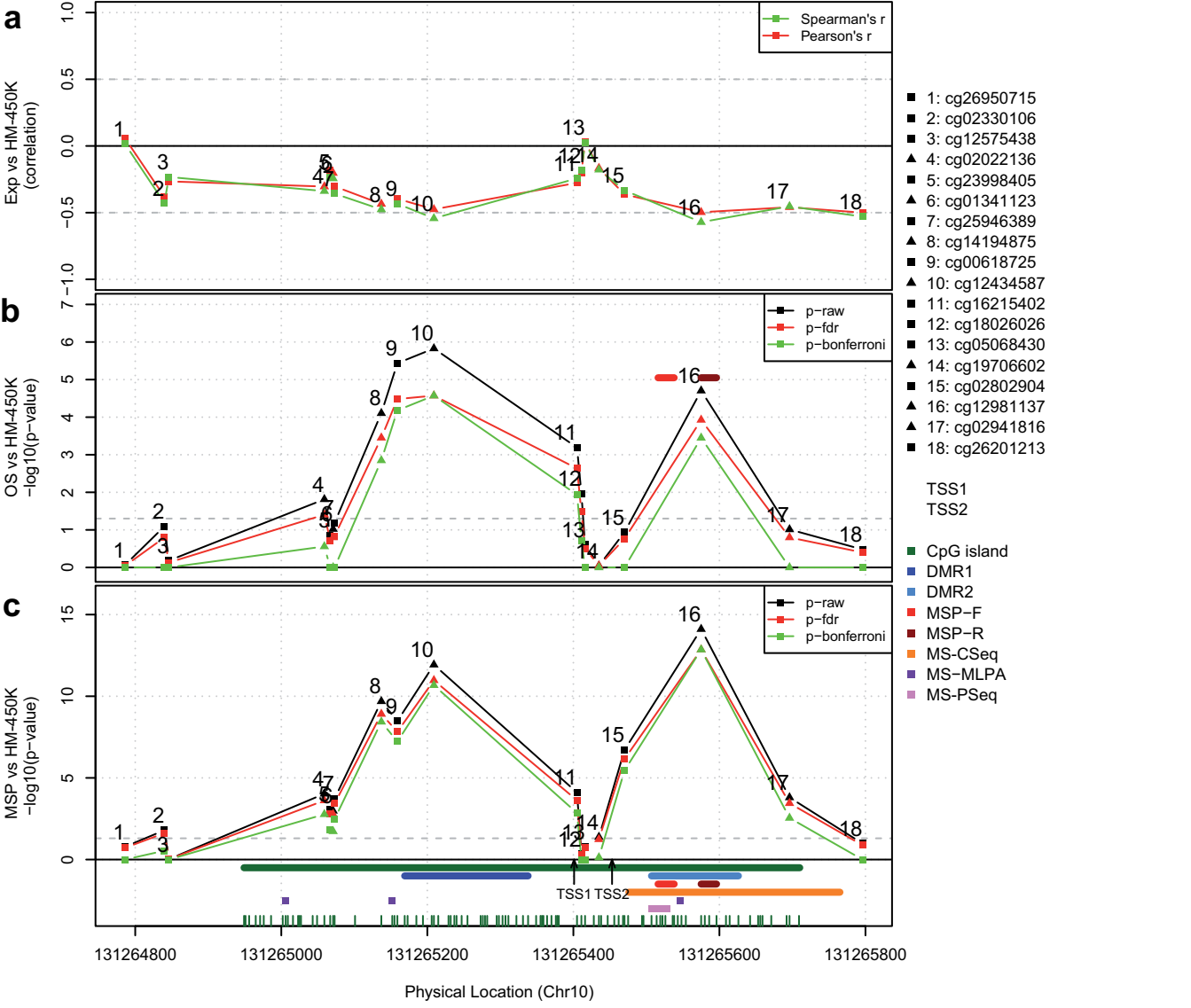
correspond to the red line and grey area, respectively. Dark green dotted lines indicate the threshold used to define methylated and unmethylated samples. (**b**) The receiver operating characteristic (ROC) curve is provided, where sensitivity (true positive rate) is plotted against 1- specificity (false positive rate). Accuracy is measured by the area under the ROC curve (AUC). Performance criteria are given for the optimal cut-off below the curve: optimal cut-off, sensitivity (sens), specificity (spec), positive predictive value (pv+), negative predictive value (pv-) and area under the curve (auc). The Kaplan-Meier curves for 58 patients are displayed for MSP-based classification (**c**) for the predicted methylation status obtained by the MGMT-STP27 model (**d**), and with additional stratification by treatment arm (RT+TMZ or RT treatment arm) (**e**) to visualize the predictive value of *MGMT* methylation for benefit from TMZ. Results of log-rank tests are given below each survival representation. M, methylated; U, unmethylated.

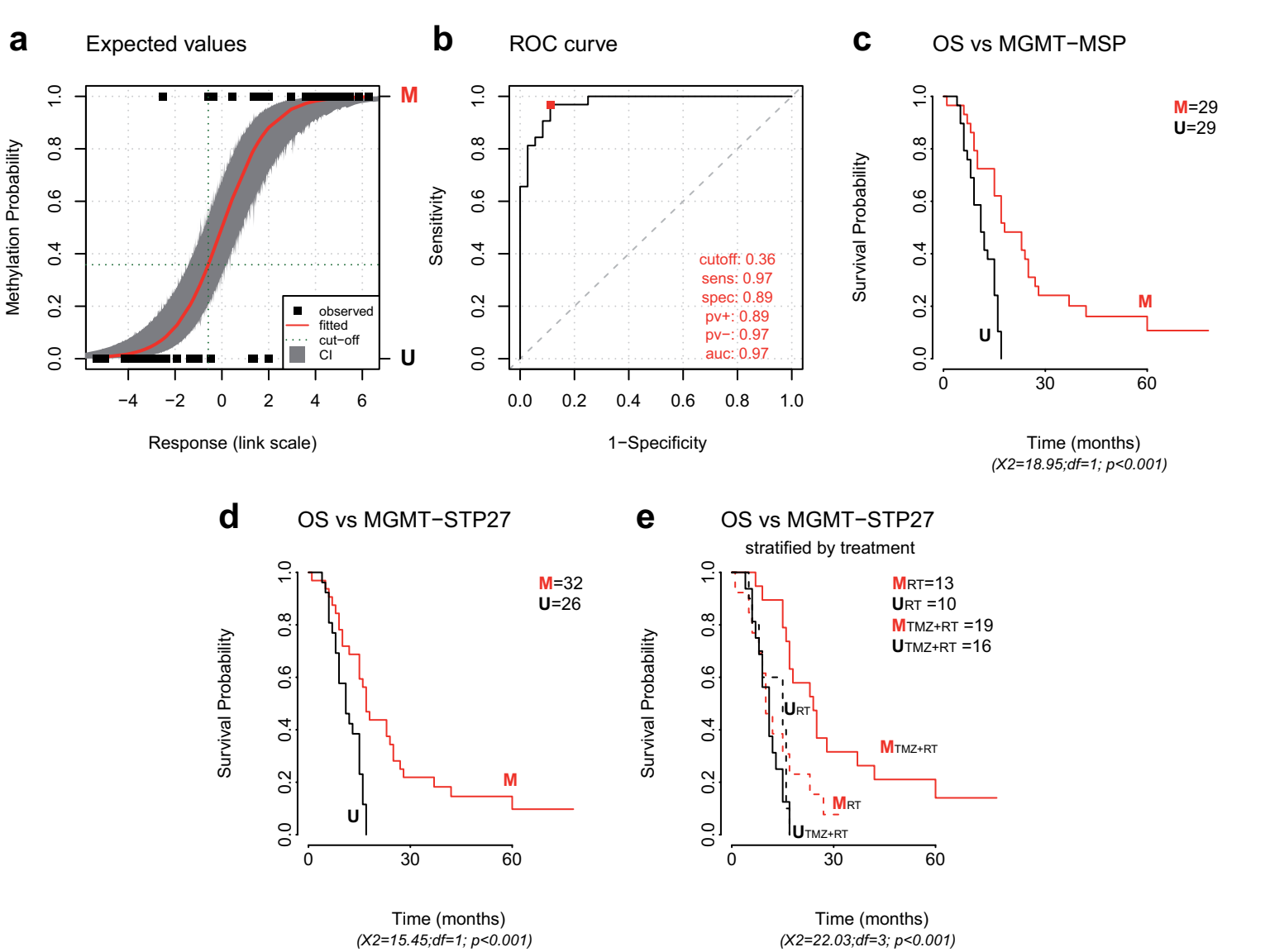
Figure 3. Validation of MGMT-STP27 in external datasets. The plots **a** and **d** represent the estimated probability of methylation against logit-transformed response fitted for the E-GBM, and VB-Glioma-III datasets using STP27. Fitted values and their prediction intervals [PI] at 95%, estimated by simulation, correspond to the red line and grey area, respectively. Dark green dotted lines indicate the threshold used to define methylated and unmethylated samples according to STP27. The observed values are visualized by full black squares, indicating same or different classification by STP27 or MS-PSeq for E-GBM and MS-MLPA test for VB-Glioma-III, respectively. The Kaplan-Meier curves are based on classification by MS-PSeq for E-GBM (**b**), the MS-MLPA test for VB-Glioma-III (**e**), and based on the respective predicted methylation status using MGMT-STP27, in (**c**) and (**f**). Results of log-rank tests are given below each survival representation. M, methylated; U, unmethylated.

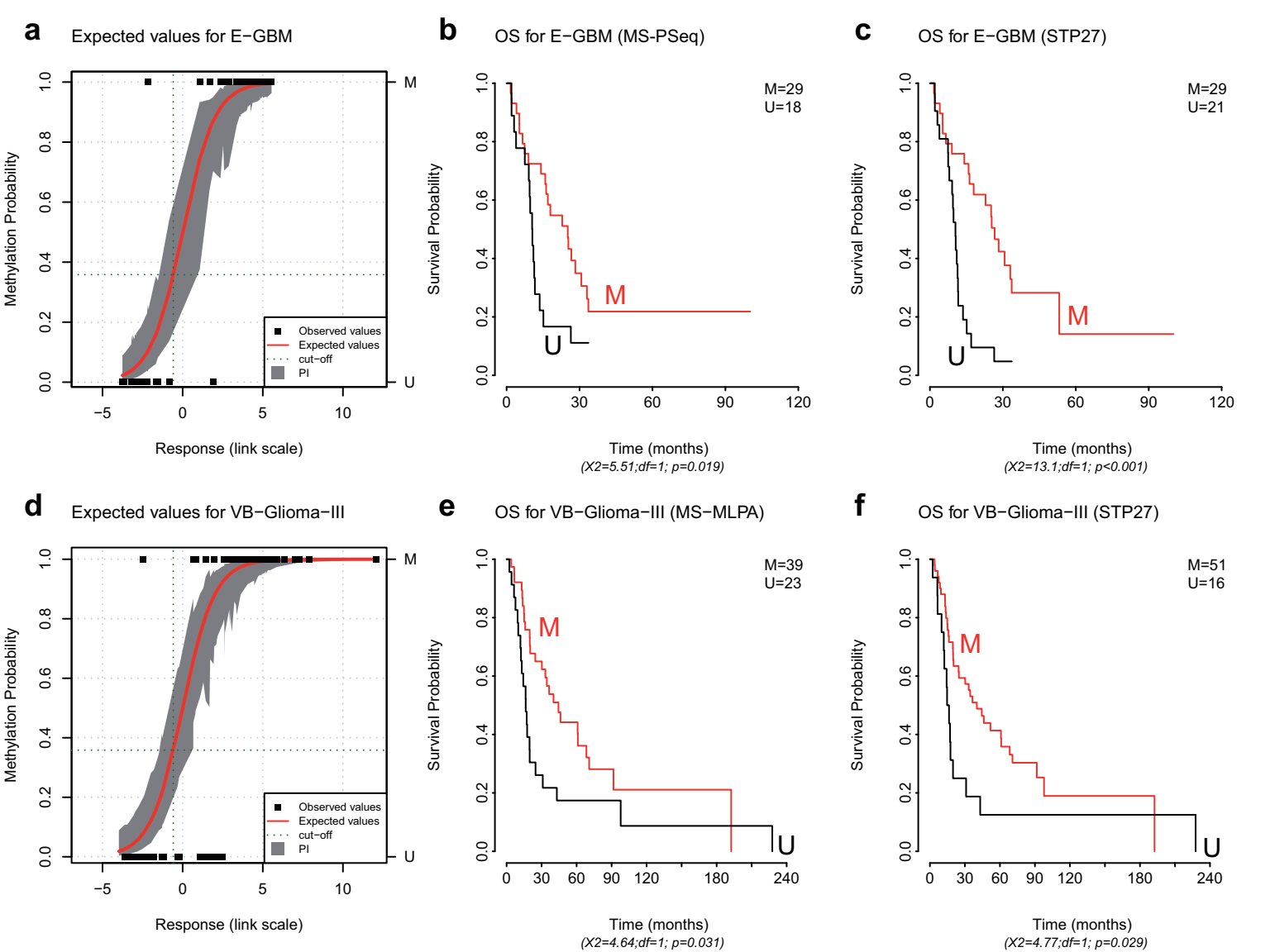
Figure 4. MGMT-STP27 based prediction in external datasets. The first plots in (**a**) and (**c**) represent the estimated values (probability of methylation fitted against response fitted in link space) for the GBM-TCGA and T-Glioma-II/III datasets. Fitted values and their prediction intervals [PI] at 95%, estimated by simulation, correspond to the red line and grey area, respectively. Dark green dotted lines indicate the threshold used to define methylated and unmethylated samples. The white squares correspond to the deduced methylation status. The Kaplan-Meier curves are based on classification by prediction using MGMT-STP27 for TCGA-GBM (**b**), T-Glioma II/III (**d**), or only T-Glioma-III (**e**), Results of log-rank tests are given below each survival representation. M, methylated; U, unmethylated.

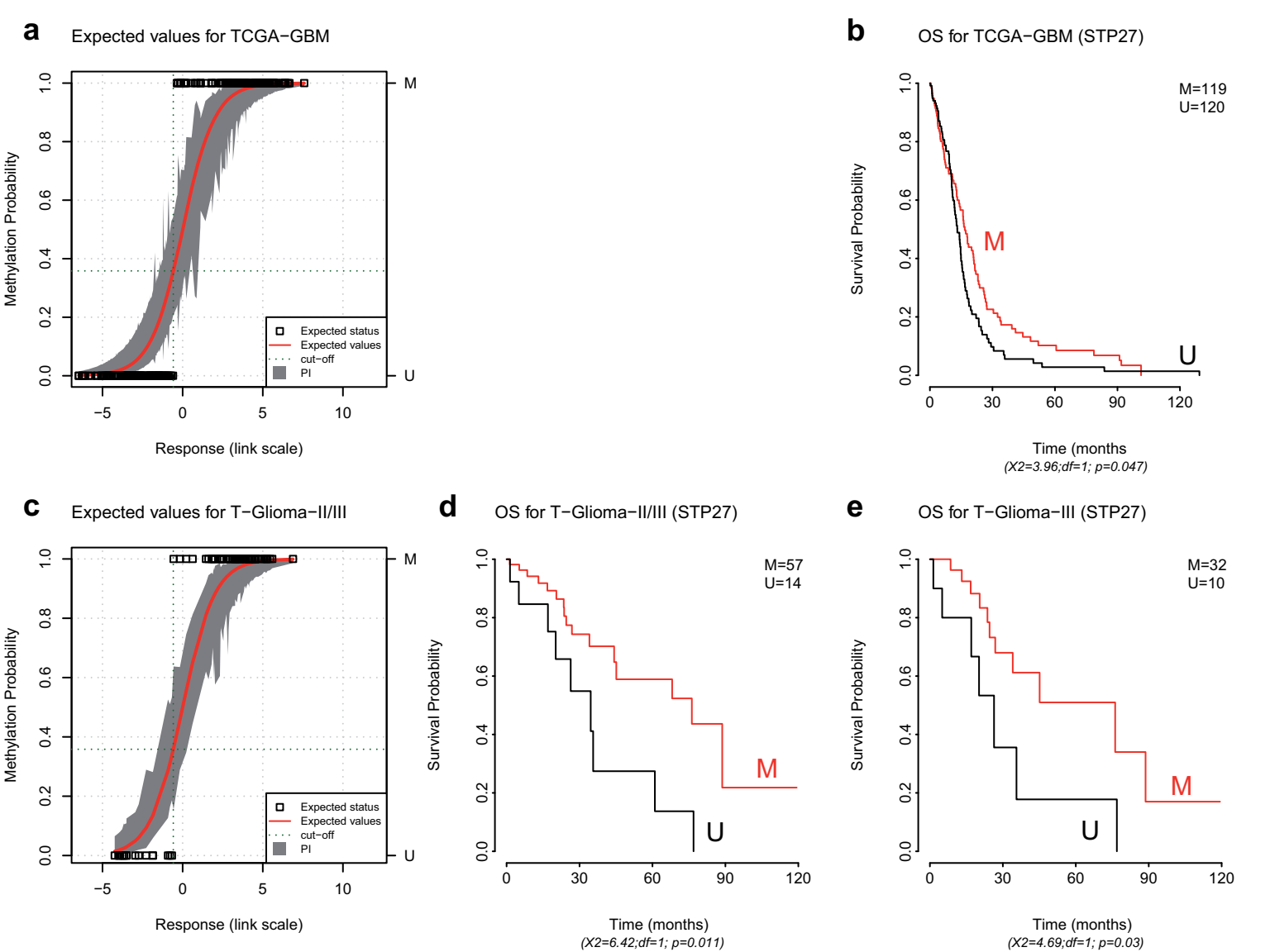
Figure 5. Distribution of *MGMT* methylation and CIMP status. The dendrogram for each dataset is provided. The five datasets were centered and normalized by probes followed by unsupervised hierarchical classifications of the 1000 most variable probes (autosomes only) using Ward's algorithm and Euclidean distance to establish CIMP classification (green rectangle for non-CIMP and red for CIMP). The methylation status of the *MGMT* promoter predicted by MGMT-STP27, blue for unmethylated, and red for methylated, is provided as label. Sample description comprise CIMP status as established in the respective original publication (if available), gender, IDH1 status (mutated or not, with additional annotation for TCGA-GBM; u, unvalidated; v, validated), classification into methylation clusters according to Noushmehr et al. [29] (cluster annotation Level 4 data, TCGA data portal), and gene expression based glioblastoma classification using a modified model from Verhaak et al. [48], tumor grade (for T-Glioma-II/III), and methylation status of *MGMT* promoter based on MSP, MS-MLPA, and MS-PSeq, unmethylated, light green; methylated, darkgreen. The color code for the labels is displayed.

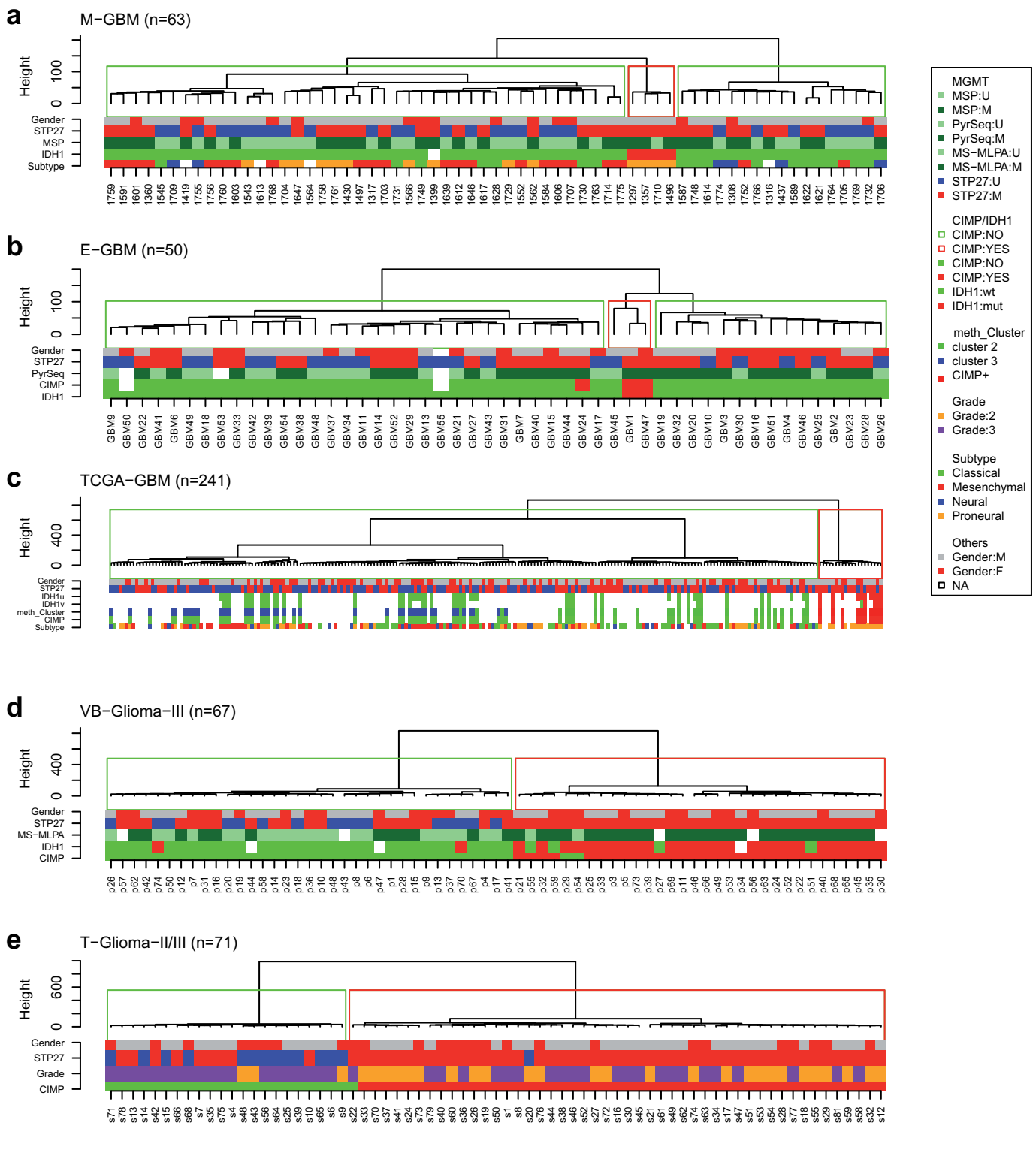
Figure 6. Proportion of predicted *MGMT* promoter methylation in CIMP+ or CIMP- gliomas. For all 5 glioma datasets the proportion of CIMP+ (**a**), the proportion of *MGMT* methylation (**b**), and the proportion of *MGMT* methylation in CIMP+ (c) and CIMP- (d) tumors, respectively, is given. The CIMP status and the *MGMT* promoter methylation status are derived from unsupervised classification and MGMT-STP27, respectively.

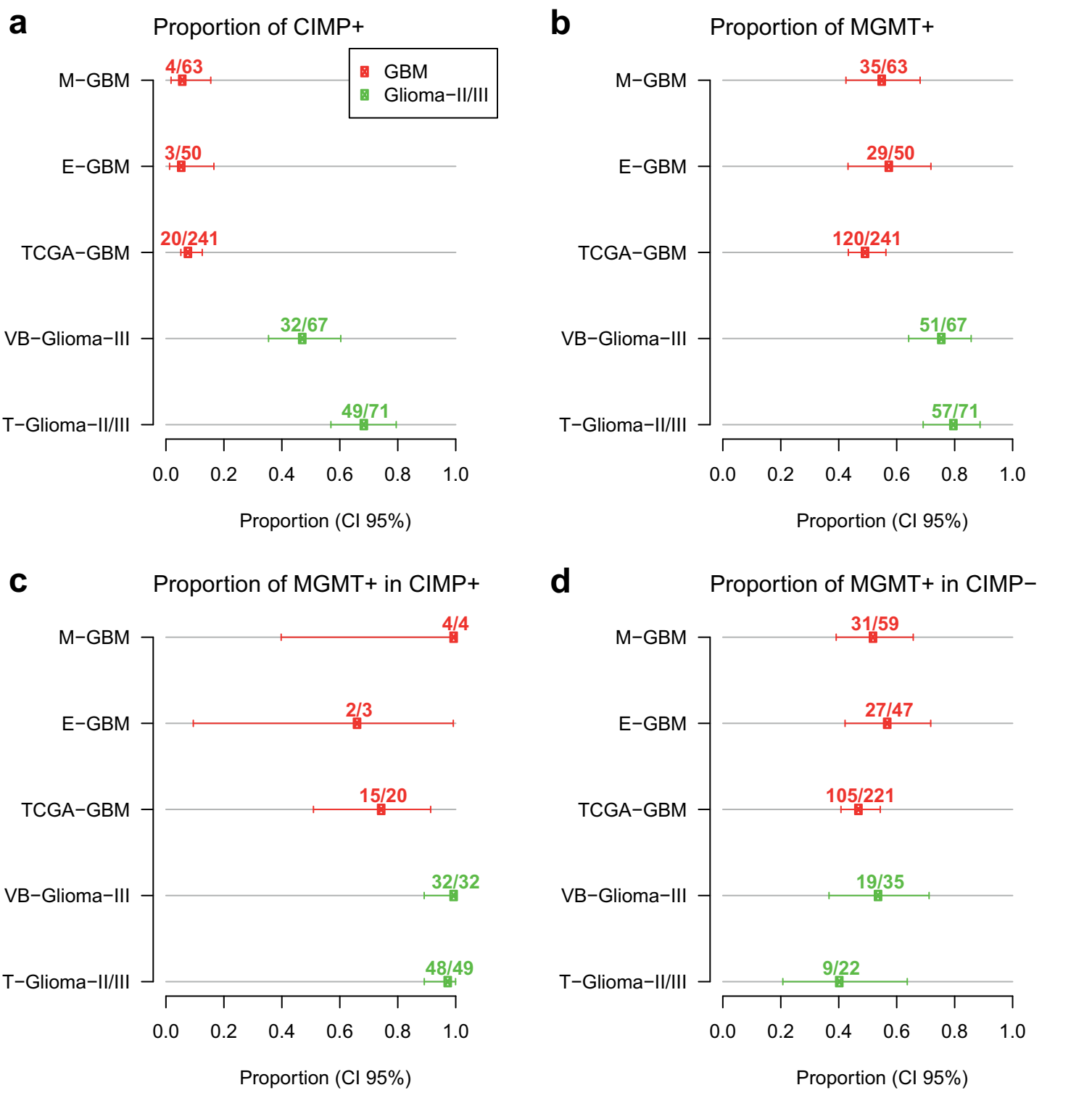

a
Exp vs HM-450K
(correlation)
Spearman's r
Pearson's r
b
OS vs HM-450K
−log10(p−value)
p−raw
p−fdr
p−bonferroni
c
MSP vs HM-450K
−log10(p−value)
p−raw
p−fdr
p−bonferroni
TSS1 TSS2
131264800
131265000
131265200
131265400
131265600
131265800
Physical Location (Chr10)
1: cg26950715
2: cg02330106
3: cg12575438
4: cg02022136
5: cg23998405
6: cg01341123
7: cg25946389
8: cg14194875
9: cg00618725
10: cg12434587
11: cg16215402
12: cg18026026
13: cg05068430
14: cg19706602
15: cg02802904
16: cg12981137
17: cg02941816
18: cg26201213
TSS1
TSS2
CpG island
DMR1
DMR2
MSP−F
MSP−R
MS-CSeq
MS−MLPA
MS-PSeq


a
Expected values
Methylation Probability
Response (link scale)
observed
fitted
cut−off
CI
M
U
b
ROC curve
Sensitivity
1−Specificity
cutoff: 0.36
sens: 0.97
spec: 0.89
pv+: 0.89
pv−: 0.97
auc: 0.97
c
OS vs MGMT−MSP
Survival Probability
Time (months)
M=29
U=29
(X2=18.95;df=1; p<0.001)
d
OS vs MGMT−STP27
Survival Probability
Time (months)
M=32
U=26
(X2=15.45;df=1; p<0.001)
e
OS vs MGMT−STP27
stratified by treatment
Survival Probability
Time (months)
MRT=13
URT =10
MTMZ+RT =19
UTMZ+RT =16
(X2=22.03;df=3; p<0.001)


a
Expected values for E−GBM
Methylation Probability
Response (link scale)
Observed values
Expected values
cut−off
PI
M
U
b
OS for E−GBM (MS-PSeq)
M=29
U=18
Survival Probability
Time (months)
(X2=5.51;df=1; p=0.019)
c
OS for E−GBM (STP27)
M=29
U=21
Survival Probability
Time (months)
(X2=13.1;df=1; p<0.001)
d
Expected values for VB−Glioma−III
Methylation Probability
Response (link scale)
Observed values
Expected values
cut−off
PI
e
OS for VB−Glioma−III (MS−MLPA)
M=39
U=23
Survival Probability
Time (months)
(X2=4.64;df=1; p=0.031)
f
OS for VB−Glioma−III (STP27)
M=51
U=16
Survival Probability
Time (months)
(X2=4.77;df=1; p=0.029)


a
Expected values for TCGA−GBM
Methylation Probability
Response (link scale)
Expected status
Expected values
cut−off
PI
M
U
b
OS for TCGA−GBM (STP27)
Survival Probability
Time (months
(X2=3.96;df=1; p=0.047)
M=119
U=120
c
Expected values for T−Glioma−II/III
d
OS for T−Glioma−II/III (STP27)
M=57
U=14
(X2=6.42;df=1; p=0.011)
e
OS for T−Glioma−III (STP27)
M=32
U=10
(X2=4.69;df=1; p=0.03)


a
M−GBM (n=63)
Height
0
100
Gender
STP27
MSP
IDH1
Subtype
b
E−GBM (n=50)
Height
0
100
Gender
STP27
PyrSeq
CIMP
IDH1
c
TCGA−GBM (n=241)
Height
0
400
Gender
STP27
IDH1u
IDH1v
meth_Cluster
CIMP
Subtype
d
VB−Glioma−III (n=67)
Height
0
400
Gender
STP27
MS−MLPA
IDH1
CIMP
e
T−Glioma−II/III (n=71)
Height
0
600
Gender
STP27
Grade
CIMP
MGMT
MSP:U
MSP:M
PyrSeq:U
PyrSeq:M
MS−MLPA:U
MS−MLPA:M
STP27:U
STP27:M
CIMP/IDH1
CIMP:NO
CIMP:YES
CIMP:NO
CIMP:YES
IDH1:wt
IDH1:mut
meth_Cluster
cluster 2
cluster 3
CIMP+
Grade
Grade:2
Grade:3
Subtype
Classical
Mesenchymal
Neural
Proneural
Others
Gender:M
Gender:F
NA


a
Proportion of CIMP+
GBM
Glioma−II/III
M−GBM
E−GBM
TCGA−GBM
VB−Glioma−III
T−Glioma−II/III
4/63
3/50
20/241
32/67
49/71
0.0
0.2
0.4
0.6
0.8
1.0
Proportion (CI 95%)
b
Proportion of MGMT+
35/63
29/50
120/241
51/67
57/71
Proportion (CI 95%)
c
Proportion of MGMT+ in CIMP+
4/4
2/3
15/20
32/32
48/49
Proportion (CI 95%)
d
Proportion of MGMT+ in CIMP−
31/59
27/47
105/221
19/35
9/22
Proportion (CI 95%)

Code TuBa	Age	Sex	OS [months]	Status	PrGBM	TMZ/RT
1076	NA	NA	NA	NA	NA	NA
1297	36	M	60	dead	yes	yes
1308	45	F	17	dead	yes	yes
1316	68	M	9	dead	no	no
1317	26	M	17	dead	yes	yes
1357	36	M	28	dead	yes	yes
1360	65	M	9	dead	yes	yes
1399	NA	F	NA	NA	NA	NA
1419	48	F	27	dead	no	no
1430	37	M	15	dead	yes	yes
1437	48	M	16	dead	yes	yes
1496	49	M	10	dead	yes	no
1497	37	M	15	dead	yes	yes
1543	27	M	11	dead	yes	yes
1545	58	M	9	dead	yes	no
1552	56	M	11	dead	yes	yes
1562	56	F	23	dead	yes	yes
1564	45	M	9	dead	no	no
1566	60	F	15	dead	yes	no
1584	56	M	17	dead	yes	yes
1587	60	F	44	alive	yes	yes
1589	48	M	8	dead	yes	yes
1591	44	M	37	dead	yes	yes
1601	55	F	24	dead	yes	yes
1603	53	M	17	dead	yes	no
1606	48	F	9	dead	yes	yes
1612	57	M	10	dead	yes	no
1613	27	M	11	dead	yes	yes
1614	49	M	42	dead	yes	yes
1617	63	M	7	dead	yes	yes
1621	62	M	25	dead	yes	yes
1622	62	M	25	dead	yes	yes
1628	49	F	15	dead	yes	yes
1639	43	M	16	dead	yes	no
1646	63	M	15	dead	yes	yes
1647	44	F	5	dead	yes	no
1703	48	M	17	dead	yes	no
1704	48	M	4	dead	yes	yes
1705	53	M	13	dead	yes	yes
1706	44	M	9	dead	yes	no
1707	61	F	12	dead	yes	yes
1709	41	M	16	dead	yes	no
1710	43	M	17	dead	yes	yes
1714	52	M	78	alive	yes	yes
1729	49	M	16	dead	yes	no
1730	49	M	15	dead	yes	no
1731	54	M	15	dead	yes	no
1732	62	F	5	dead	yes	no
1748	59	M	6	dead	yes	no

1749	52	F	33	alive	yes	no
1752	54	M	23	dead	yes	no
1755	45	F	6	dead	yes	no
1756	56	M	8	dead	yes	no
1758	37	M	18	dead	yes	yes
1759	53	M	1	dead	yes	no
1760	63	M	8	dead	yes	no
1761	51	M	27	dead	yes	no
1763	62	M	12	dead	yes	no
1764	32	M	7	dead	yes	yes
1766	57	M	6	dead	yes	yes
1768	55	F	6	dead	yes	yes
1769	52	M	9	dead	yes	yes
1774	64	M	25	dead	no	no
1775	52	M	78	alive	yes	yes
1780	NA	NA	NA	NA	NA	NA
2900	NA	NA	NA	NA	NA	NA
2901	NA	NA	NA	NA	NA	NA
2903	NA	NA	NA	NA	NA	NA

"Normal" Brain Tissue	PatientID	MSP-MGMT	IDH1-status	CIMP
yes	NB	U	NA	NA
no	P-206	M	mut	yes
no	P-209	M	wt	no
no	P-210	M	wt	no
no	P-216	U	wt	no
no	P-221	M	mut	yes
no	P-222	M	wt	no
no	1399	M	NA	no
no	P-236	M	wt	no
no	P-241	M	wt	no
no	P-240	U	wt	no
no	55	M	mut	yes
no	P-241	M	wt	no
no	106	U	wt	no
no	94	U	wt	no
no	236	U	wt	no
no	P-112	M	wt	no
no	P-114	U	wt	no
no	159	U	wt	no
no	344	U	wt	no
no	90	M	wt	no
no	157	U	wt	no
no	288	M	wt	no
no	360	M	wt	no
no	323	M	wt	no
no	274	U	wt	no
no	428	M	wt	no
no	106	U	wt	no
no	396	M	wt	no
no	444	M	wt	no
no	P-113	M	wt	no
no	P-113	M	wt	no
no	458	U	wt	no
no	566	U	wt	no
no	461	U	wt	no
no	574	U	wt	no
no	63	M	wt	no
no	129	U	wt	no
no	161	U	wt	no
no	211	M	wt	no
no	222	U	wt	no
no	390	U	wt	no
no	547	U	mut	yes
no	57	M	wt	no
no	83	U	wt	no
no	147	M	wt	no
no	242	U	wt	no
no	97	U	wt	no
no	345	M	wt	no

no	527	M	wt	no
no	246	M	wt	no
no	15	U	wt	no
no	283	M	wt	no
no	12	M	wt	no
no	84	M	wt	no
no	254	U	wt	no
no	284	M	wt	no
no	328	U	wt	no
no	356	U	wt	no
no	506	U	wt	no
no	558	U	wt	no
no	563	U	wt	no
no	100	U	wt	no
no	57	M	wt	no
yes	NB (rep2)	U	NA	NA
yes	NB_BNT_92	U	NA	NA
yes	NB_BNT_104	U	NA	NA
yes	NB_BNT_148	U	NA	NA

Expression Subtype	Trial	STP27link	STP27response	STP27class	cg00618725
NA	NA	-1.35414492	0.20519355	U	-3.42229711
Proneural	II	5.24997476	0.99477974	M	1.09732403
Neural	II	-0.56761824	0.36178658	M	-1.74116
NA	II	-0.58297811	0.35824762	M	-3.17887999
Mesenchymal	II	-3.92717878	0.01931861	U	-2.76826385
Proneural	II	3.8658949	0.97948549	M	1.43413797
Mesenchymal	II	1.67811008	0.84265411	M	-3.27528806
Proneural	NA	3.80175551	0.97815627	M	0.22313066
NA	II	-0.38279168	0.40545376	M	-1.53197982
Proneural	II	5.34180624	0.99523559	M	0.02092065
Neural	II	-0.46572198	0.38562929	M	-3.02381719
Proneural	III	4.36968733	0.98750296	M	-0.23950915
Mesenchymal	II	5.06517912	0.99372682	M	-3.10937265
Proneural	III	-3.97314631	0.01846671	U	-3.50787221
Classical	III	-3.38013919	0.03292196	U	-3.09603937
Mesenchymal	III	-3.97373131	0.01845611	U	-3.01859928
Proneural	II	4.6035714	0.99008332	M	-1.7111505
NA	II	-3.31871702	0.03493464	U	-1.30458223
Proneural	III	-2.72701148	0.06139816	U	-2.72102712
Mesenchymal	III	-5.34611391	0.00474403	U	-3.53740425
Classical	III	4.43887174	0.98832858	M	1.32309795
Classical	III	-3.27375056	0.03648276	U	-3.16452472
Mesenchymal	III	3.67024901	0.97516249	M	-0.51676433
Mesenchymal	III	2.96603324	0.95101582	M	-1.82660502
Mesenchymal	III	-2.51025461	0.07514241	U	-2.82740581
Mesenchymal	III	-4.99719455	0.00671153	U	-3.04320664
Mesenchymal	III	1.99918294	0.88071127	M	-0.72867908
NA	III	-3.89425067	0.01995242	U	-3.28360627
Classical	III	3.60044961	0.97341464	M	-3.1260154
Classical	III	4.81300059	0.99194201	M	0.79930677
Classical	II	5.86416602	0.99716865	M	0.31064419
Classical	II	4.26071507	0.98608418	M	-0.03815013
Classical	III	-2.38167052	0.08458113	U	-3.04587993
Mesenchymal	III	-4.08917623	0.01647699	U	-3.05960318
Mesenchymal	III	-3.8032384	0.02181207	U	-3.09379595
Proneural	III	2.01144136	0.88199312	M	0.94539006
Mesenchymal	III	1.50081594	0.81769614	M	-1.68493689
Proneural	III	-3.474633	0.03004268	U	-3.30444607
Mesenchymal	III	-3.47940737	0.02990387	U	-3.04351556
Neural	III	5.14543137	0.99420778	M	0.4963276
Mesenchymal	III	-3.07583028	0.04411532	U	-2.82630259
Neural	III	-3.71848851	0.02369552	U	-3.34218339
Proneural	III	1.2900384	0.78415369	M	-2.26418065
Mesenchymal	III	3.80623294	0.97825173	M	-0.10378648
Proneural	III	-3.85952422	0.02064291	U	-3.44856061
Classical	III	1.45699613	0.81107281	M	-0.53574595
Classical	III	-1.01876004	0.265269	U	-2.37698207
Classical	III	-3.38097638	0.03289532	U	-2.91817331
Classical	III	1.37098234	0.79753882	M	-2.70558527

Classical	III	3.71915267	0.97631984	M	0.24825754
Mesenchymal	III	4.13618464	0.98426774	M	-1.01399215
Neural	III	-3.0801818	0.04393218	U	-3.74179709
Mesenchymal	III	1.45901301	0.81138167	M	-2.39745462
Proneural	III	3.64845712	0.97462917	M	-1.73590387
Mesenchymal	III	0.43837652	0.60787212	M	-3.20495147
Mesenchymal	III	-3.19147156	0.03948793	U	-3.59023392
Proneural	III	6.26194375	0.9980961	M	1.90540162
Classical	III	1.37339975	0.79792888	M	-2.58225903
Classical	III	-3.04611858	0.04538534	U	-2.90630495
Classical	III	-4.1393187	0.0156838	U	-3.28995631
Mesenchymal	III	-3.57926522	0.02713911	U	-2.87748481
Classical	III	-1.91280337	0.12866624	U	-2.79563828
Neural	III	-2.95286775	0.04960115	U	-2.9346134
Mesenchymal	III	5.46171148	0.99577167	M	1.25706206
NA	NA	-3.97902062	0.01836053	U	-3.32179977
NA	NA	-2.79232517	0.05774032	U	-3.226305
NA	NA	-3.40822602	0.03203937	U	-3.65328891
NA	NA	-1.31857271	0.21105586	U	-3.04142861

cg01341123	cg02022136	cg02330106	cg02802904	cg02941816	cg05068430	cg12434587
-1.24331302	-2.30658285	1.25765198	-6.08330775	-1.98029317	-6.26007944	-4.72584593
-3.29276153	-5.27569998	0.38374859	-5.46267489	-2.14376015	-5.0530342	0.8801676
-1.55347613	-2.25804525	2.15926015	-5.54735569	0.45155452	-5.0565698	-2.88327136
-3.00928189	-4.06119049	1.31436374	-5.29848306	-1.90856991	-4.63636699	-5.64469229
-2.68034904	-4.25589508	1.3498265	-6.34986568	-3.39160967	-4.54073952	-5.51347848
1.52117568	1.74125473	2.98219195	-0.77330803	2.01452877	-4.62765511	-0.55078194
-1.81958057	-3.1495966	2.11624936	-4.52279102	0.87567247	-4.26019185	-4.7766114
-0.48891446	-0.2013222	1.7653049	-5.16273532	-2.2989416	-4.30882176	-1.60886747
-1.39089607	-1.74188071	1.52759921	-5.37647083	-1.91661813	-6.84232896	-1.69844285
0.66269592	-0.86970213	2.48428783	-3.24863698	-3.08379968	-4.12085594	0.78390984
-1.39243616	-1.35928734	1.94295968	-5.20876695	0.03270205	-4.72990393	-4.22557123
2.06062209	2.62305567	3.21996788	-0.33640398	1.15278292	-4.13973297	-0.48006868
-2.61214846	-2.70348091	2.22149148	-3.4032979	-2.99450742	-6.47806689	0.57256061
-2.28961864	-3.32663649	1.62488604	-5.39088103	-3.41774554	-5.85810598	-4.8603404
-1.29438737	-3.75806011	1.12591178	-5.73425982	-1.37676031	-4.33608997	-5.19882168
-2.78362138	-5.02298112	0.37922822	-5.64665871	-3.46791565	-4.68833735	-5.0786142
-2.12505211	-3.52625885	2.58168104	-3.64367367	-1.81424622	-3.33835064	0.97145088
-0.5151677	-1.05202434	2.70672342	-5.28633905	-2.49664028	-4.53580163	-4.5191994
-2.61831908	-2.37823261	0.6766963	-5.77373897	2.1372357	-4.93795925	-4.45593484
-3.06026448	-4.95135737	0.07482011	-6.26861763	-3.66765279	-4.55205349	-5.96475641
-2.31664975	-0.06011547	2.87748022	-2.91371827	1.16690383	-4.37695212	-0.46379062
-2.80084896	-4.5985601	0.04615434	-5.52534347	-2.74540005	-6.56872561	-5.17307405
-0.13473334	-1.173555	1.26906965	-5.49568126	-2.99551341	-4.94069205	-0.55745197
-1.06658661	-0.8961834	1.57124511	-1.79234705	-0.92070316	-3.45804491	-0.08918209
-1.8536612	-3.98228517	1.67573846	-5.36066011	-0.77911172	-4.64121929	-4.42792432
-1.52723544	-2.88290136	2.02897361	-7.6374845	-3.17441523	-6.51343446	-6.03689235
-2.07166526	-2.5273367	1.5657519	-5.96523181	-1.62636232	-4.32457598	-4.00343573
-1.6885261	-2.62343665	1.00485209	-5.40229514	-3.06162809	-4.29482258	-5.43853533
-1.56701405	-1.85726247	2.39985373	-5.57576225	1.32303093	-6.24140943	-2.87737035
-1.70189119	-3.93609404	2.27224026	-5.44048771	1.20524389	-4.87099788	0.32522339
0.63388802	1.6399516	2.5658426	-1.25334383	-1.81052649	-3.99074051	1.30021158
0.24558127	-0.31446718	1.16750926	-2.43492351	-2.07254447	-4.43018561	0.12978624
-2.52393693	-3.25643439	2.95958814	-5.26576313	-2.10516251	-5.53557447	-4.80575946
-1.03294512	-1.88177925	1.83243097	-5.39936368	-2.06439461	-4.69232267	-5.89419909
-3.16911327	-4.58221643	0.04293473	-4.99448864	-3.68330445	-4.20761563	-5.31487334
0.59947138	0.54312237	1.83727934	-1.9828368	-0.85620181	-4.86400371	0.31352256
1.68598499	2.09802318	2.79159443	-3.65186409	-2.64964398	-4.88301689	-1.27956776
-2.88025733	-4.65085846	-0.65864294	-5.72515546	-3.19870744	-4.61806698	-5.16850152
1.50113948	1.20847235	2.78548391	-5.50627697	-3.44825269	-4.54770216	-5.20857816
0.9588049	1.2793946	2.40781018	-1.76187934	1.37386702	-3.60779922	0.48479812
-3.07740918	-4.51862803	1.49364459	-4.92044565	-2.60640521	-3.8694754	-5.00597624
-3.0477176	-5.0540624	0.715179	-6.49809065	-3.71657749	-6.52029776	-5.09506504
-2.26943975	-1.74501252	2.43904423	-5.13294825	-0.1479255	-5.06970439	-3.27633123
0.3438824	-0.37653153	2.01698791	-5.24782851	-2.318437	-4.18840299	-0.37021338
-2.27323243	-3.7553132	2.34393327	-5.18211083	-2.88617569	-4.30368645	-5.27120362
-3.1736748	-4.59486824	2.73941312	-4.98663878	-3.21697598	-4.18181022	-3.441229
-2.29304192	-4.08676468	2.53662617	-4.64566196	-2.50382003	-4.63945735	-4.06823364
-2.61083081	-4.70184558	1.56459801	-5.41837477	-2.86418614	-4.80530736	-5.4050908
1.05878582	1.05865296	0.94294678	-4.22942023	-2.30938695	-5.4117573	-2.67067722

0.61623893	-0.05101571	1.9642317	-2.17574382	-3.24979653	-4.8598276	0.0389083
0.12727275	-0.40520566	1.37807606	-5.36185856	-0.69537326	-4.77212469	-0.08041088
-1.10178916	-1.833752	2.06539323	-5.5851723	-2.17319528	-4.38879608	-5.70856635
-2.08330295	-2.87057028	1.89031889	-2.14879613	-1.1192648	-6.1874374	-0.8891419
1.41940611	1.28083995	0.34941493	-3.35387863	1.24327254	-4.38212886	-2.58479
0.23338479	-0.37101448	0.57101169	-3.0164527	-1.84858173	-6.01378904	-4.9350931
-2.17328639	-4.05718086	1.61233555	-5.4406775	-2.94948193	-4.70709406	-5.34470588
2.17354109	1.88866229	3.07190811	-3.72148326	2.02718011	-5.99306449	1.44112755
1.44864394	-0.51888033	2.62655741	-5.79339013	-3.11312054	-4.78152468	-4.05472219
1.27116322	0.74159674	2.39524334	-5.3530726	-3.23114583	-5.42034881	-5.39932986
-2.69039119	-4.46442374	0.39340987	-5.46704247	-1.82275831	-5.05565816	-5.04891497
-2.25283913	-3.74256149	0.42560546	-5.40839219	-3.30479958	-6.68552944	-5.11760048
-2.25606201	-3.11739554	1.88523909	-5.42044403	-1.28733317	-4.61335863	-5.03279757
-2.18380148	-3.99211375	2.42955738	-5.39213106	-2.50930041	-4.72240192	-5.25351163
1.16852439	1.44168417	2.32930134	-4.41774335	-0.46862872	-4.87495643	0.94069012
-1.86524666	-2.87933988	0.82356006	-6.41515318	-2.50805452	-5.17257945	-5.5649878
-1.64096101	-2.86475781	1.18371381	-6.0971581	-2.32234186	-5.9608646	-5.7932301
-1.48490151	-2.56174209	1.25007025	-5.58345891	-2.4886035	-5.28996	-5.84544738
-1.25850068	-2.52886794	1.22011678	-5.49222368	-1.84842574	-6.03176775	-4.88154898

cg12575438	cg12981137	cg14194875	cg16215402	cg18026026	cg19706602	cg23998405
2.94063939	-3.43748691	-3.52578691	-2.90542426	-5.45235981	-4.59545205	-3.4884984
0.67237709	0.5013628	0.96742175	-2.60439879	-4.13163652	-3.5549326	-8.37116203
2.78018779	-3.63684587	0.15036375	-2.82642165	-5.40717538	-3.89506119	-3.73460751
3.25771097	-2.08233417	-3.59469919	-2.8800804	-5.41766502	-3.99499353	-5.56134411
2.37362904	-5.76664893	-3.80614056	-2.96730784	-6.16477803	-3.69047907	-5.41690306
3.72895904	-0.17845603	1.44504504	-2.28608612	-4.7445545	-4.24765435	1.43677475
2.58409593	-0.13565042	-3.04565956	-2.93531687	-6.64898933	-3.90048877	-4.26786004
2.07364244	0.35430347	0.65796208	-2.81135131	-4.47816818	-4.15719458	-3.49036678
1.5931547	-4.11144764	-2.26351185	-3.05277513	-4.80780881	-3.76994085	-2.81164915
2.51768114	0.65524899	-0.50374591	-2.22728261	-4.58800296	-3.56478462	-1.81695205
2.79773451	-2.76316916	-2.85283986	-2.72327477	-4.09798264	-3.90872337	-2.62095125
3.53818964	0.32509537	-0.07110419	-2.10604375	-4.67590044	-2.28465722	2.24906215
1.84418865	0.47690899	-1.76032407	-2.73974511	-4.16641995	-3.29293251	-5.84133361
2.50326922	-6.18786361	-3.74683415	-2.93187379	-4.90879389	-3.78294793	-5.05138408
2.34414583	-5.35520738	-3.93842061	-2.95350838	-4.55588267	-3.96388971	-4.43406082
1.55121051	-6.06430972	-4.27191579	-2.73209584	-3.98829714	-3.83471853	-5.84583403
3.02356944	-0.2482867	0.05056125	-2.37933733	-5.21262854	-3.40829735	-6.23799638
3.36408777	-5.67557581	-0.54167761	-2.9720465	-5.2008265	-4.07722833	-3.57045874
2.85438026	-5.07289491	-2.12876826	-2.58617813	-4.2636354	-3.40670249	-5.30298653
2.25662846	-7.04146703	-4.39941871	-3.0613683	-6.05509857	-4.64606191	-6.23399242
2.69335094	0.39051033	1.19286725	-1.97334075	-2.89519827	-3.41828511	-4.93559768
2.42645639	-5.25502259	-4.04584461	-2.57388397	-4.90469591	-3.10205431	-5.27876078
2.0544692	-0.38583732	-1.86608378	-2.72522039	-5.99449606	-4.05411926	-2.0548677
2.086651	-1.41237221	-0.5774011	-2.2489289	-3.05757405	-3.71796054	-2.95540339
2.51514751	-4.85486838	-3.34869985	-2.75184395	-3.92118662	-4.13624188	-5.28236929
3.26219668	-6.6238094	-3.8980644	-3.10280383	-5.30329755	-4.06807633	-4.16605869
2.33553039	-0.22898377	-0.17987072	-3.0321477	-4.8368577	-3.51006363	-4.44105941
2.16192415	-5.77374537	-3.42151824	-2.88922954	-4.30444709	-3.9232206	-3.86347938
2.94966442	0.85872373	-2.45228513	-2.86036776	-4.20311336	-3.82094948	-1.87197193
2.23595487	0.34543341	0.66638296	-2.92872734	-4.97554542	-3.77985019	-4.99324623
3.07809592	0.9253272	-0.165096	-2.49160936	-4.07948478	-2.79192908	0.04217573
1.94313153	-0.13949968	-0.09206429	-2.55369586	-4.31795044	-3.53948988	-0.92216661
3.43469559	-4.50111073	-3.03426683	-2.527247	-4.59672221	-3.37645589	-4.81339384
2.7525316	-5.72489737	-3.49234094	-2.4676176	-4.66317133	-3.94108502	-2.92437802
1.32533687	-5.74586521	-4.00755954	-2.75606396	-4.3330902	-3.64587774	-6.25455109
2.01915469	-2.67185504	0.91238442	-2.75099102	-4.59415555	-4.37739463	-0.20306521
3.23042947	-2.31663622	0.90179666	-2.33356344	-3.73799572	-3.67862681	1.4490455
-0.57535784	-5.47445249	-3.93115166	-2.77992554	-6.19498409	-4.01183326	-5.39313619
3.44113633	-5.45680453	-3.0770877	-2.7001804	-3.55570753	-3.44694563	-0.32061693
2.82188894	0.61346347	1.16135771	-2.35578418	-3.72139059	-2.98218543	0.42421938
2.73388181	-5.13646062	-4.06877828	-2.52664949	-4.19508209	-3.09326379	-5.93914837
0.66460965	-5.77944629	-3.69121449	-3.19939191	-4.85232085	-4.07962513	-6.36243669
3.14263496	-1.40810079	-0.63658925	-2.70483684	-5.36441778	-3.16842683	-5.11142493
2.90791579	-0.34558712	-0.32270805	-2.01657422	-3.56686308	-3.65561904	-1.05333097
3.34427364	-5.83146455	-3.32731817	-2.71318542	-3.71304277	-4.45622919	-5.09799858
2.81043185	-1.13407262	1.42082953	-2.60000545	-4.62199062	-3.19636451	-5.84900597
3.21558724	-3.4496232	-0.33726151	-2.93416856	-5.06411598	-3.70242943	-5.65807051
3.18613361	-5.23875567	-4.07230189	-2.913198	-5.18413434	-3.00642195	-5.4723922
3.44820987	-1.66531409	-2.28798783	-2.8709439	-6.00756069	-3.86445049	0.41581834

2.48409622	-0.67234665	-1.08136635	-2.69775916	-5.39026522	-3.56898379	-2.34136624
2.16619852	-0.1543251	-2.26505404	-2.77486552	-5.7205482	-3.89216909	-1.66124008
3.24618955	-4.74142391	-3.28586035	-3.31037581	-4.85724706	-3.4112344	-3.12224108
2.32378567	-2.58388747	-0.83689265	-2.70368745	-4.96182627	-4.08350437	-4.95096225
2.28588178	0.74408053	-1.9702878	-2.65831666	-4.90146182	-3.52780408	0.63199825
1.20199976	-1.38362635	-0.58162112	-2.88103623	-6.60725706	-3.65360634	-1.12523356
2.3880506	-5.06856362	-4.0461551	-2.8571423	-6.00439438	-4.07601063	-5.3848478
3.70668013	1.27450719	1.94112029	-2.46263234	-4.94171091	-3.05302158	0.76993712
3.08919947	-0.87526224	-3.02546087	-2.79545906	-4.62032765	-3.77828715	-1.74865555
3.02652778	-4.88059467	-2.88762087	-2.88346891	-5.54049374	-3.43356373	-0.88963608
2.3167151	-6.25993852	-4.39321899	-2.64564657	-4.76744508	-3.67856187	-5.4486569
1.33283507	-5.61635015	-4.42281623	-2.81170257	-6.46364576	-3.75123279	-5.07603589
3.27303396	-3.86585343	-3.91423724	-2.76782656	-4.51488406	-3.30858656	-4.6445519
3.61751483	-4.86290389	-4.02219692	-3.01745358	-3.58586957	-3.28147438	-5.63117345
2.37131021	0.69547321	1.4209702	-2.71216757	-4.37096749	-3.56988193	0.76665608
2.15220023	-5.79330008	-3.99742866	-3.1826955	-5.08113128	-4.96988646	-4.0138151
2.53451581	-4.38254853	-4.02884559	-2.98949671	-4.6835968	-4.0250375	-4.14876079
2.10487698	-5.01763062	-3.57749339	-3.05970762	-5.03481603	-4.40589632	-3.18849772
2.41455455	-3.31050485	-3.26433807	-3.44278983	-4.5700721	-4.7746709	-3.40151905

cg25946389	cg26201213	cg26950715
-1.31539602	0.37808312	2.67231699
-3.18590307	0.14387418	0.22858595
0.37786826	1.17990638	2.31792686
-2.41381142	1.26503045	3.10006892
-2.22983247	-0.78563742	2.75928456
1.98829618	2.29960912	2.99044665
-1.04122034	1.4036828	2.99951142
0.50982262	1.34911124	2.555347
-0.76605975	0.27622057	1.86100107
1.66115235	1.51127673	2.49217958
0.08727932	1.21532413	2.62436285
2.59785201	2.77829221	2.98240956
-2.15052644	1.1235314	0.24059181
-1.64585632	-0.62654785	2.92256779
-0.45425832	0.48949995	2.7793284
-2.57619041	0.26905685	1.90730778
-1.69421395	-1.27767504	1.78920691
-0.23220507	1.70670687	3.0212675
-0.77391098	2.23344448	2.64047554
-2.74537386	0.93875095	2.15604699
-1.23151856	2.37866421	2.45890708
-1.87839769	1.31808287	2.10579103
-0.19704359	0.35120901	2.0986126
0.0392484	0.22241532	2.72841408
-1.97555809	0.55570136	2.29555458
-1.23549407	1.1522079	2.97736721
-1.47000427	0.7926917	1.39062143
-1.29733312	0.15703587	2.10307872
1.64637059	2.02960727	2.77022907
-1.98816962	1.11153676	1.26876049
0.87672255	2.22569378	2.57197575
0.29032767	0.43780317	2.601111
-1.73043558	1.89784503	3.20495338
-1.2820553	0.50266061	2.76791135
-2.88777728	-0.7491149	1.25206166
1.08005833	1.13724675	0.9719961
2.39896527	-0.00910962	2.72220672
-2.70159929	-0.44805929	2.36239768
2.05937794	1.66117858	2.77142745
1.42856278	1.48522522	1.87439837
-2.61094673	1.59368862	2.7372781
-2.77796506	-1.01979562	3.03537953
-1.69263944	1.78168208	2.47273265
0.38470503	0.96635088	2.54051367
-1.69807386	1.44854122	3.06987849
-2.4904372	2.14750507	-0.27888976
-2.58115433	0.30813909	2.26104889
-2.40742904	1.70265218	2.98546761
1.26332694	1.93157373	3.2807547

0.74143282	0.7708461	2.08878668
0.14347137	0.25099092	2.49997646
-0.54432052	0.5919029	2.87505356
-0.59573066	1.07350817	2.58878666
1.65475319	1.72364451	2.7547123
0.31514194	0.21161705	2.40772502
-2.17876654	0.2047283	2.97749289
2.45743296	2.23261736	3.07678801
1.78510514	1.99865983	1.40287631
1.3633976	1.1182408	2.66364419
-2.64050063	0.52977147	2.76354677
-2.12628886	-0.88936657	1.73754487
-1.89915282	1.80006355	2.56636696
-1.92114044	2.22426048	3.41641987
1.47911315	1.93628826	3.06556202
-1.91221796	0.54001165	1.82314228
-1.54424868	0.66278214	2.35522452
-1.33413421	0.73118425	2.38546181
-1.10894277	0.88429417	2.52858859

Table S2. MGMT Prediction Models for M-GBM dataset

Probe_ID	Code_Fig1	MAP_INFO	Present in 27K	Design type (450K)	rpearson
cg26950715	1	131264786	no	II	0.057
cg02330106	2	131264840	yes	II	-0.381
cg12575438	3	131264846	no	II	-0.266
cg02022136	4	131265059	no	I	-0.304
cg23998405	5	131265067	no	I	-0.180
cg01341123	6	131265071	no	II	-0.203
cg25946389	7	131265073	yes	II	-0.303
cg14194875	8	131265137	no	I	-0.435
cg00618725	9	131265159	no	I	-0.396
cg12434587	10	131265209	yes	I	-0.475
cg16215402	11	131265405	no	I	-0.276
cg18026026	12	131265411	no	I	-0.203
cg05068430	13	131265416	no	I	0.031
cg19706602	14	131265435	yes	I	-0.167
cg02802904	15	131265470	no	I	-0.363
cg12981137	16	131265575	yes	I	-0.497
cg02941816	17	131265696	yes	II	-0.459
cg26201213	18	131265796	yes	II	-0.500
Models					
STP27	-	-	-	-	-
STP450	-	-	-	-	-

see Supplement for description of abbreviations and Illumina webpage for description of probes and d

rspearman	p	AIC	AICc	BIC	CovPat	MVIF	CutOff	Sens	Spec
0.015	1	96.099	96.159	100.538	0.000	-	0.445	0.656	0.556
-0.429	1	92.228	92.289	96.667	0.000	-	0.392	0.844	0.444
-0.233	1	98.029	98.090	102.468	0.000	-	0.475	0.125	0.972
-0.338	1	82.781	82.842	87.220	0.000	-	0.509	0.625	0.833
-0.233	1	86.894	86.954	91.333	0.015	-	0.518	0.594	0.833
-0.244	1	87.279	87.339	91.718	0.000	-	0.569	0.531	0.889
-0.358	1	84.081	84.141	88.520	0.000	-	0.553	0.625	0.861
-0.476	1	57.606	57.667	62.045	0.029	-	0.377	0.906	0.861
-0.433	1	63.047	63.108	67.486	0.015	-	0.599	0.750	0.944
-0.543	1	47.511	47.571	51.950	0.029	-	0.355	0.875	0.944
-0.243	1	82.450	82.511	86.889	0.000	-	0.338	0.875	0.500
-0.184	1	97.474	97.535	101.913	0.000	-	0.489	0.406	0.750
0.026	1	96.156	96.217	100.595	0.000	-	0.523	0.438	0.833
-0.178	1	93.883	93.944	98.322	0.000	-	0.422	0.813	0.472
-0.333	1	70.956	71.017	75.396	0.015	-	0.609	0.594	0.972
-0.571	1	37.644	37.704	42.083	0.029	-	0.545	0.906	0.944
-0.453	1	83.758	83.819	88.197	0.015	-	0.387	0.813	0.667
-0.530	1	95.390	95.451	99.829	0.000	-	0.576	0.281	0.917
-	2	36.143	36.327	42.801	0.029	1.160	0.358	0.969	0.889
-	4	36.621	37.256	47.718	0.029	8.884	0.509	0.938	0.944

esign ty

Kappa	Accuracy
0.210	0.603
0.281	0.632
0.102	0.574
0.463	0.735
0.432	0.721
0.428	0.721
0.492	0.750
0.765	0.882
0.702	0.853
0.822	0.912
0.366	0.676
0.159	0.588
0.277	0.647
0.278	0.632
0.578	0.794
0.852	0.926
0.474	0.735
0.205	0.618
0.853	0.926
0.882	0.941

Table S4. Annotation of TCGA-GBM dataset

bcr_patient_barcode	dataset	Subtype	CIMP	MGMT-STP27 (response)
TCGA-19-0964	HM-27K	Classical	0	0.998330419
TCGA-06-2565	HM-27K	NA	0	0.997486082
TCGA-06-0877	HM-27K	Classical	0	0.908505291
TCGA-02-0007	HM-27K	Classical	0	0.278962809
TCGA-02-0009	HM-27K	Classical	0	0.538506655
TCGA-02-0021	HM-27K	Classical	0	0.754682945
TCGA-06-0122	HM-27K	Mesenchymal	0	0.156252119
TCGA-12-1098	HM-27K	Classical	0	0.001862951
TCGA-12-3648	HM-27K	Classical	0	0.682345526
TCGA-14-0783	HM-27K	Mesenchymal	0	0.995492886
TCGA-14-0817	HM-27K	Neural	0	0.007156551
TCGA-14-1396	HM-27K	Mesenchymal	0	0.994618723
TCGA-14-1402	HM-27K	NA	0	0.981814064
TCGA-16-1062	HM-27K	Classical	0	0.751231943
TCGA-19-1789	HM-27K	Classical	0	0.997113054
TCGA-27-1833	HM-27K	Classical	0	0.02863833
TCGA-28-1750	HM-27K	Mesenchymal	0	0.058666463
TCGA-19-1390	HM-27K	Proneural	0	0.998021697
TCGA-14-0790	HM-27K	NA	0	0.979065038
TCGA-12-1598	HM-27K	Proneural	0	0.996178567
TCGA-06-2563	HM-27K	NA	0	0.995770084
TCGA-06-0125	HM-27K	Classical	0	0.998725868
TCGA-12-3646	HM-27K	Proneural	0	0.996766016
TCGA-19-0957	HM-27K	Proneural	0	0.26276641
TCGA-28-1746	HM-27K	Proneural	0	0.996587035
TCGA-02-0037	HM-27K	Classical	0	0.179789127
TCGA-02-0083	HM-27K	Neural	0	0.998238686
TCGA-06-0876	HM-27K	Classical	0	0.994714849
TCGA-28-2509	HM-27K	NA	0	0.99144555
TCGA-32-1982	HM-27K	NA	0	0.990831092
TCGA-16-1060	HM-27K	Mesenchymal	0	0.055697552
TCGA-14-0786	HM-27K	Classical	0	0.972938253
TCGA-06-0155	HM-27K	Mesenchymal	0	0.016997011
TCGA-14-1829	HM-27K	NA	0	0.009230552
TCGA-28-2514	HM-27K	NA	0	0.009377393
TCGA-12-0821	HM-27K	Neural	0	0.01386001
TCGA-14-1455	HM-27K	Proneural	0	0.958310211
TCGA-26-1440	HM-27K	Classical	0	0.049201815
TCGA-16-0861	HM-27K	Mesenchymal	0	0.107165507
TCGA-02-0113	HM-27K	Neural	0	0.410497982
TCGA-06-2566	HM-27K	NA	0	0.293244442
TCGA-02-0027	HM-27K	Classical	0	0.527775567
TCGA-02-0069	HM-27K	Proneural	0	0.970083936
TCGA-06-0875	HM-27K	Proneural	0	0.040611377
TCGA-06-1801	HM-27K	Proneural	0	0.175264197
TCGA-06-2558	HM-27K	NA	0	0.009365492
TCGA-12-0826	HM-27K	Classical	0	0.225857217
TCGA-12-1088	HM-27K	Mesenchymal	0	0.065267535

TCGA-12-3653	HM-27K	Classical	0	0.039718838
TCGA-15-1447	HM-27K	Proneural	0	0.992564055
TCGA-19-0960	HM-27K	Proneural	0	0.998604282
TCGA-27-1836	HM-27K	NA	0	0.973357516
TCGA-27-1838	HM-27K	NA	0	0.181813527
TCGA-28-1755	HM-27K	NA	0	0.004661451
TCGA-14-1459	HM-27K	NA	0	0.941975859
TCGA-14-1454	HM-27K	Proneural	0	0.010033461
TCGA-12-3644	HM-27K	Neural	0	0.968177473
TCGA-12-1099	HM-27K	Proneural	0	0.025634894
TCGA-02-0001	HM-27K	Mesenchymal	0	0.116894242
TCGA-02-0006	HM-27K	Mesenchymal	0	0.354295582
TCGA-02-0074	HM-27K	Proneural	0	0.23886496
TCGA-02-0085	HM-27K	Mesenchymal	0	0.990359033
TCGA-02-0086	HM-27K	Mesenchymal	0	0.523972007
TCGA-06-0147	HM-27K	Mesenchymal	0	0.949905212
TCGA-06-2561	HM-27K	NA	0	0.069567473
TCGA-12-1093	HM-27K	Mesenchymal	0	0.001526901
TCGA-12-1095	HM-27K	Mesenchymal	0	0.002243917
TCGA-12-1599	HM-27K	Neural	0	0.863103528
TCGA-14-1034	HM-27K	Mesenchymal	0	0.935653571
TCGA-14-1037	HM-27K	Mesenchymal	0	0.072660405
TCGA-14-2554	HM-27K	NA	0	0.102677937
TCGA-14-2555	HM-27K	Classical	0	0.972693898
TCGA-19-0962	HM-27K	Mesenchymal	0	0.993504187
TCGA-19-1791	HM-27K	Classical	0	0.034568769
TCGA-26-1443	HM-27K	Classical	0	0.983663742
TCGA-27-1835	HM-27K	NA	0	0.951404907
TCGA-28-1751	HM-27K	Mesenchymal	0	0.037376984
TCGA-28-2506	HM-27K	NA	0	0.948647263
TCGA-28-2513	HM-27K	NA	0	0.009338676
TCGA-32-2616	HM-27K	Mesenchymal	0	0.965870228
TCGA-28-1752	HM-27K	Neural	0	0.989987363
TCGA-27-2526	HM-27K	NA	0	0.011694456
TCGA-27-1832	HM-27K	Mesenchymal	0	0.029674946
TCGA-19-2625	HM-27K	Mesenchymal	0	0.048536409
TCGA-19-1786	HM-27K	Classical	0	0.002849194
TCGA-02-0060	HM-27K	Proneural	0	0.975086192
TCGA-02-0057	HM-27K	Mesenchymal	0	0.129337636
TCGA-02-0055	HM-27K	Mesenchymal	0	0.218941228
TCGA-02-0054	HM-27K	Mesenchymal	0	0.916127822
TCGA-02-0038	HM-27K	Classical	0	0.164440024
TCGA-02-0003	HM-27K	Proneural	0	0.115882884
TCGA-02-0033	HM-27K	Mesenchymal	0	0.991631635
TCGA-02-0034	HM-27K	Mesenchymal	0	0.183508551
TCGA-06-1086	HM-27K	Mesenchymal	0	0.006302393
TCGA-06-2569	HM-27K	NA	0	0.007729294
TCGA-12-0822	HM-27K	Mesenchymal	0	0.703784561
TCGA-12-1089	HM-27K	Classical	0	0.038041334
TCGA-12-1097	HM-27K	Neural	0	0.00471193
TCGA-14-0865	HM-27K	Proneural	0	0.914291794

TCGA-14-3477	HM-27K	Proneural	0	0.015370591
TCGA-16-1055	HM-27K	Mesenchymal	0	0.01723816
TCGA-19-1388	HM-27K	Mesenchymal	0	0.976336377
TCGA-19-1389	HM-27K	Mesenchymal	0	0.166154656
TCGA-27-1830	HM-27K	Mesenchymal	0	0.467201477
TCGA-27-1831	HM-27K	NA	0	0.03478327
TCGA-28-1757	HM-27K	Classical	0	0.934500058
TCGA-32-1986	HM-27K	NA	0	0.026688194
TCGA-32-2615	HM-27K	Mesenchymal	0	0.049972179
TCGA-41-2571	HM-27K	Proneural	0	0.06706634
TCGA-28-1747	HM-27K	NA	0	0.920882681
TCGA-28-1745	HM-27K	Mesenchymal	0	0.267675593
TCGA-27-2527	HM-27K	NA	0	0.860753193
TCGA-27-2519	HM-27K	NA	0	0.006198627
TCGA-27-1834	HM-27K	Neural	0	0.972997501
TCGA-26-1438	HM-27K	Mesenchymal	0	0.054068739
TCGA-19-2624	HM-27K	Proneural	0	0.036458851
TCGA-19-1385	HM-27K	Mesenchymal	0	0.989580984
TCGA-19-0955	HM-27K	Classical	0	0.474639973
TCGA-14-1827	HM-27K	Neural	0	0.954473406
TCGA-14-1451	HM-27K	Proneural	0	0.002043454
TCGA-14-0871	HM-27K	Proneural	0	0.00332835
TCGA-14-0736	HM-27K	NA	0	0.039167731
TCGA-12-1096	HM-27K	Mesenchymal	0	0.016444267
TCGA-12-1090	HM-27K	Classical	0	0.995328493
TCGA-06-2567	HM-27K	NA	0	0.99157017
TCGA-06-2564	HM-27K	NA	0	0.06098418
TCGA-06-1087	HM-27K	Proneural	0	0.026764717
TCGA-06-1084	HM-27K	Mesenchymal	0	0.937397752
TCGA-06-0882	HM-27K	Neural	0	0.00842615
TCGA-06-0881	HM-27K	Mesenchymal	0	0.006849411
TCGA-06-0878	HM-27K	Mesenchymal	0	0.023738864
TCGA-06-0141	HM-27K	Mesenchymal	0	0.123520667
TCGA-06-0139	HM-27K	Mesenchymal	0	0.136731892
TCGA-06-0130	HM-27K	Mesenchymal	0	0.156141827
TCGA-02-2470	HM-27K	NA	0	0.015536742
TCGA-02-0115	HM-27K	Neural	0	0.964247289
TCGA-02-0107	HM-27K	Mesenchymal	0	0.112107864
TCGA-02-0052	HM-27K	Neural	0	0.331725028
TCGA-02-0047	HM-27K	Proneural	0	0.118340395
TCGA-02-0024	HM-27K	Proneural	0	0.105293533
TCGA-02-0011	HM-27K	Proneural	0	0.965415245
TCGA-02-2466	HM-27K	NA	0	0.032211745
TCGA-06-0143	HM-27K	Mesenchymal	0	0.113077986
TCGA-06-1800	HM-27K	Mesenchymal	0	0.994141188
TCGA-12-1091	HM-27K	Classical	0	0.969705326
TCGA-12-1092	HM-27K	Mesenchymal	0	0.925135819
TCGA-14-1825	HM-27K	Proneural	0	0.001534259
TCGA-19-0963	HM-27K	Neural	0	0.150853644
TCGA-32-1976	HM-27K	NA	0	0.991114157
TCGA-41-2573	HM-27K	Proneural	0	0.990443232

TCGA-41-2575	HM-27K	Proneural	0	0.019756425
TCGA-27-1837	HM-27K	NA	0	0.989016549
TCGA-14-1401	HM-27K	Proneural	0	0.957618369
TCGA-14-0813	HM-27K	Proneural	0	0.994084289
TCGA-12-3651	HM-27K	Neural	0	0.989974577
TCGA-12-1602	HM-27K	Proneural	0	0.994455645
TCGA-06-2559	HM-27K	NA	0	0.996494286
TCGA-02-0046	HM-27K	Proneural	0	0.996478078
TCGA-02-0043	HM-27K	Classical	0	0.999492021
TCGA-02-0064	HM-27K	Mesenchymal	0	0.991298598
TCGA-02-0071	HM-27K	Mesenchymal	0	0.251844354
TCGA-02-0075	HM-27K	Mesenchymal	0	0.994161069
TCGA-02-0099	HM-27K	Mesenchymal	0	0.996412984
TCGA-06-0124	HM-27K	Mesenchymal	0	0.992616714
TCGA-06-0133	HM-27K	Neural	0	0.203097319
TCGA-06-0879	HM-27K	Classical	0	0.991814815
TCGA-12-0828	HM-27K	Neural	0	0.428052248
TCGA-12-0829	HM-27K	Mesenchymal	0	0.974353776
TCGA-12-1094	HM-27K	Classical	0	0.010609353
TCGA-14-0867	HM-27K	Mesenchymal	0	0.996494219
TCGA-14-1794	HM-27K	Proneural	0	0.035993564
TCGA-16-0848	HM-27K	Proneural	0	0.97892215
TCGA-16-1063	HM-27K	NA	0	0.048338267
TCGA-19-2621	HM-27K	Classical	0	0.99300517
TCGA-26-1799	HM-27K	Neural	0	0.088293165
TCGA-28-1753	HM-27K	NA	0	0.024318304
TCGA-32-1970	HM-27K	NA	0	0.015561859
TCGA-32-2632	HM-27K	Mesenchymal	0	0.071135652
TCGA-41-3392	HM-27K	Neural	0	0.060745459
TCGA-28-2502	HM-27K	NA	0	0.07056148
TCGA-28-1760	HM-27K	Mesenchymal	0	0.050825763
TCGA-19-2623	HM-27K	Mesenchymal	0	0.957940907
TCGA-19-1386	HM-27K	Classical	0	0.04241051
TCGA-16-1056	HM-27K	Classical	0	0.078627608
TCGA-15-1446	HM-27K	Classical	0	0.023796181
TCGA-14-0866	HM-27K	Neural	0	0.994101497
TCGA-14-0789	HM-27K	Mesenchymal	0	0.986369689
TCGA-12-3652	HM-27K	Classical	0	0.041552682
TCGA-12-3650	HM-27K	Neural	0	0.192357093
TCGA-12-0820	HM-27K	Classical	0	0.306882803
TCGA-06-0169	HM-27K	Mesenchymal	0	0.32151847
TCGA-06-0148	HM-27K	NA	0	0.186509872
TCGA-02-0089	HM-27K	Neural	0	0.418865701
TCGA-02-2485	HM-27K	NA	0	0.007420642
TCGA-02-2486	HM-27K	NA	0	0.06724384
TCGA-06-0126	HM-27K	Classical	0	0.184693523
TCGA-12-0670	HM-27K	Neural	0	0.753622856
TCGA-14-0787	HM-27K	Classical	0	0.992587256
TCGA-14-0812	HM-27K	Neural	0	0.884922336
TCGA-14-1452	HM-27K	Mesenchymal	0	0.980145793
TCGA-14-1795	HM-27K	Proneural	0	0.989878983

TCGA-16-0846	HM-27K	Proneural	0	0.995150403
TCGA-19-1387	HM-27K	Proneural	0	0.993378926
TCGA-19-1787	HM-27K	Mesenchymal	0	0.995594055
TCGA-19-2620	HM-27K	Classical	0	0.989413262
TCGA-27-2518	HM-27K	NA	0	0.164002834
TCGA-27-2524	HM-27K	NA	0	0.017420817
TCGA-27-2528	HM-27K	NA	0	0.635996922
TCGA-32-2634	HM-27K	Proneural	0	0.996721382
TCGA-32-2638	HM-27K	Classical	0	0.979207886
TCGA-28-1749	HM-27K	Classical	0	0.127855734
TCGA-27-2523	HM-27K	NA	0	0.993759561
TCGA-19-1392	HM-27K	Proneural	0	0.986221824
TCGA-14-1453	HM-27K	Classical	0	0.011306363
TCGA-12-3649	HM-27K	Neural	0	0.022314196
TCGA-12-1600	HM-27K	Classical	0	0.061901421
TCGA-06-2562	HM-27K	NA	0	0.02872698
TCGA-06-2557	HM-27K	NA	0	0.178002532
TCGA-06-1802	HM-27K	Mesenchymal	0	0.989937375
TCGA-02-0116	HM-27K	Mesenchymal	0	0.241091106
TCGA-02-0102	HM-27K	Classical	0	0.19332825
TCGA-02-0014	HM-27K	Proneural	1	0.117710694
TCGA-02-0028	HM-27K	Proneural	1	0.998335898
TCGA-06-1805	HM-27K	Proneural	1	0.224773517
TCGA-06-2570	HM-27K	NA	1	0.980971043
TCGA-12-0827	HM-27K	Mesenchymal	1	0.805783854
TCGA-14-1456	HM-27K	NA	1	0.556609715
TCGA-14-1458	HM-27K	Proneural	1	0.958064164
TCGA-19-1788	HM-27K	Proneural	1	0.983781382
TCGA-19-2629	HM-27K	Proneural	1	0.049523516
TCGA-28-1756	HM-27K	NA	1	0.989170307
TCGA-27-2521	HM-27K	NA	1	0.998397621
TCGA-16-0849	HM-27K	Proneural	1	0.463153085
TCGA-14-1821	HM-27K	Proneural	1	0.943672808
TCGA-06-0129	HM-27K	Proneural	1	0.977648165
TCGA-06-0128	HM-27K	Proneural	1	0.977943137
TCGA-02-2483	HM-27K	NA	1	0.966409576
TCGA-02-0080	HM-27K	Proneural	1	0.85690658
TCGA-02-0010	HM-27K	Proneural	1	0.183573273
TCGA-16-0850	HM-27K	Proneural	1	0.078993973
TCGA-02-0114	HM-27K	Proneural	1	0.993850031
TCGA-06-5414	HM-450K	Classical	NA	0.026411471
TCGA-06-5415	HM-450K	Classical	NA	0.011341718
TCGA-06-5416	HM-450K	Proneural	NA	0.13975203
TCGA-06-5417	HM-450K	Proneural	NA	0.976177539
TCGA-06-5418	HM-450K	NA	NA	0.0861036
TCGA-12-5295	HM-450K	Neural	NA	0.992547423
TCGA-12-5299	HM-450K	Classical	NA	0.057375593
TCGA-12-5301	HM-450K	Neural	NA	0.992653037
TCGA-26-5132	HM-450K	Classical	NA	0.994449034
TCGA-26-5133	HM-450K	Proneural	NA	0.033480542
TCGA-26-5134	HM-450K	Proneural	NA	0.028343968

TCGA-26-5135	HM-450K	Proneural	NA	0.910096282
TCGA-26-5136	HM-450K	Mesenchymal	NA	0.942378446
TCGA-26-5139	HM-450K	Mesenchymal	NA	0.051294466
TCGA-28-5204	HM-450K	Neural	NA	0.035255743
TCGA-28-5207	HM-450K	NA	NA	0.033872393
TCGA-28-5208	HM-450K	NA	NA	0.995944053
TCGA-28-5209	HM-450K	Mesenchymal	NA	0.997280274
TCGA-28-5213	HM-450K	NA	NA	0.016157843
TCGA-28-5214	HM-450K	Mesenchymal	NA	0.244304771
TCGA-28-5215	HM-450K	Mesenchymal	NA	0.868875196
TCGA-28-5216	HM-450K	Mesenchymal	NA	0.034636178
TCGA-28-5218	HM-450K	Mesenchymal	NA	0.022101363
TCGA-28-5219	HM-450K	Classical	NA	0.966226487
TCGA-28-5220	HM-450K	Classical	NA	0.025217251
TCGA-32-5222	HM-450K	Proneural	NA	0.992933499
TCGA-76-4925	HM-450K	Proneural	NA	0.998237477
TCGA-76-4926	HM-450K	Mesenchymal	NA	0.018142488
TCGA-76-4927	HM-450K	Neural	NA	0.035468836
TCGA-76-4928	HM-450K	Classical	NA	0.74628518
TCGA-76-4929	HM-450K	Neural	NA	0.996332934
TCGA-76-4931	HM-450K	Classical	NA	0.022155511
TCGA-76-4932	HM-450K	Proneural	NA	0.971405986
TCGA-76-4934	HM-450K	Proneural	NA	0.897547642
TCGA-76-4935	HM-450K	Proneural	NA	0.927768584
TCGA-06-0650	HM-450K	Mesenchymal	NA	0.094094779
TCGA-06-1804	HM-450K	Classical	NA	0.98942752
TCGA-06-5408	HM-450K	Classical	NA	0.029608161
TCGA-06-5410	HM-450K	Mesenchymal	NA	0.672662967
TCGA-06-5411	HM-450K	Neural	NA	0.152043491
TCGA-06-5412	HM-450K	Mesenchymal	NA	0.964140825
TCGA-06-5413	HM-450K	Neural	NA	0.020156637
TCGA-06-5856	HM-450K	Proneural	NA	0.222756545
TCGA-06-5858	HM-450K	Mesenchymal	NA	0.452235411
TCGA-06-5859	HM-450K	Neural	NA	0.014145626
TCGA-06-6389	HM-450K	Proneural	NA	0.957375937
TCGA-06-6390	HM-450K	Classical	NA	0.039438945
TCGA-06-6391	HM-450K	Proneural	NA	0.46199731
TCGA-14-0781	HM-450K	Mesenchymal	NA	0.144585406
TCGA-15-1444	HM-450K	Proneural	NA	0.923945149
TCGA-19-5947	HM-450K	Mesenchymal	NA	0.054430349
TCGA-19-5950	HM-450K	Classical	NA	0.987087377
TCGA-19-5951	HM-450K	Classical	NA	0.136777673
TCGA-19-5952	HM-450K	Classical	NA	0.026630345
TCGA-19-5954	HM-450K	Classical	NA	0.654578442
TCGA-19-5955	HM-450K	Mesenchymal	NA	0.742994446
TCGA-19-5956	HM-450K	Proneural	NA	0.226888673
TCGA-19-5958	HM-450K	Classical	NA	0.048196133
TCGA-19-5959	HM-450K	Classical	NA	0.994068631
TCGA-19-5960	HM-450K	Proneural	NA	0.143723814
TCGA-26-1442	HM-450K	Proneural	NA	0.898414934
TCGA-28-2501	HM-450K	Mesenchymal	NA	0.961299912

TCGA-28-2510	HM-450K	Neural	NA	0.076260518
TCGA-28-6450	HM-450K	Classical	NA	0.16611129
TCGA-32-1979	HM-450K	Neural	NA	0.467259512
TCGA-32-1980	HM-450K	Neural	NA	0.036993892
TCGA-41-5651	HM-450K	Proneural	NA	0.99187475
TCGA-76-6191	HM-450K	Proneural	NA	0.021235479
TCGA-76-6192	HM-450K	Proneural	NA	0.040673492
TCGA-76-6193	HM-450K	Mesenchymal	NA	0.029230222
TCGA-76-6282	HM-450K	Mesenchymal	NA	0.102776239
TCGA-76-6285	HM-450K	Proneural	NA	0.218056739
TCGA-81-5910	HM-450K	Classical	NA	0.029113593
TCGA-87-5896	HM-450K	Classical	NA	0.105310671

MGMT-STP27 (class)
M
M
M
U
M
M
U
U
M
M
U
M
M
M
M
U
U
M
M
M
M
M
M
U
M
U
M
M
M
M
U
M
U
U
U
U
M
U
U
M
U
M
M
U
U
U
U
U

U
M
M
M
U
U
M
U
M
U
U
U
U
M
M
M
U
U
U
M
M
U
U
M
M
U
M
M
U
M
U
M
M
U
U
U
U
M
U
U
U
M
U
U
U
M
U
U
U
M
U
M

U
U
M
U
M
U
M
U
U
U
M
U
M
U
M
U
U
M
M
M
U
U
U
U
M
M
U
U
M
U
U
U
U
U
U
U
M
U
U
U
U
U
M
U
U
M
M
M
U
U
M
M

U
M
M
M
M
M
M
M
M
M
M
U
M
M
M
M
U
M
M
M
U
M
U
M
U
M
U
M
U
U
U
U
U
U
U
U
M
U
U
U
M
M
U
U
U
U
U
U
M
U
U
U
U
M
M
M
M
M

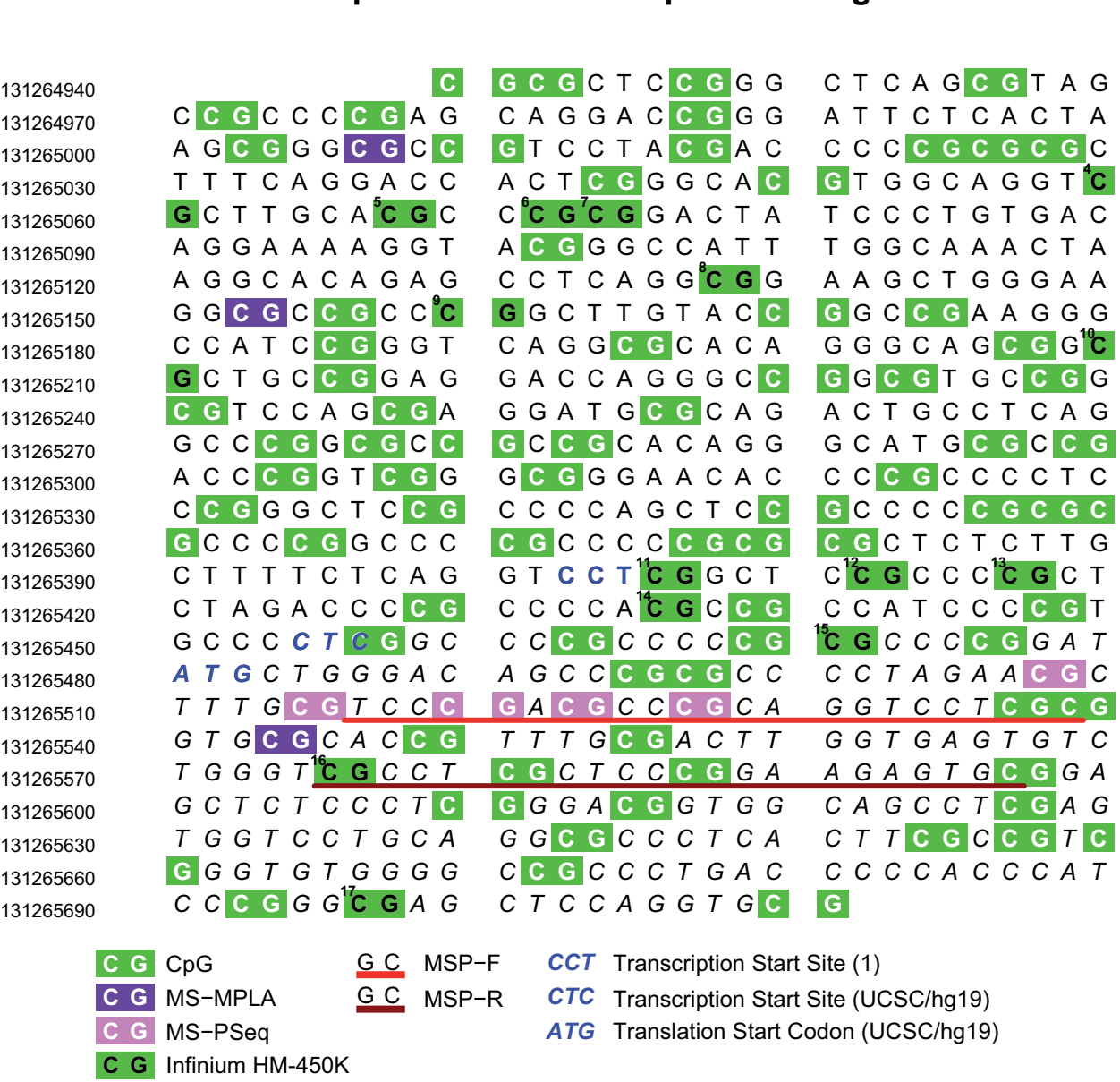
M
M
M
M
U
U
M
M
M
U
M
M
U
U
U
U
U
M
U
U
U
M
U
M
M
M
M
M
U
M
M
M
M
M
M
M
M
M
U
U
M

U
U
U
M
U
M
U
M
U
M
M
U
U

M
M
U
U
U
M
M
U
U
M
U
U
M
U
M
M
U
U
M
M
U
M
M
M
U
M
U
M
U
M
U
U
M
U
M
U
M
U
M
U
M
U
U
M
M
U
U
M
U
M
M

U
U
M
U
M
U
U
U
U
U
U
U
U

Figure S1


CpG island in MGMT promoter region
131264940
131264970
131265000
131265030
131265060
131265090
131265120
131265150
131265180
131265210
131265240
131265270
131265300
131265330
131265360
131265390
131265420
131265450
131265480
131265510
131265540
131265570
131265600
131265630
131265660
131265690
CpG
MS−MPLA
MS−PSeq
Infinium HM-450K
MSP−F
MSP−R
CCT Transcription Start Site (1)
CTC Transcription Start Site (UCSC/hg19)
ATG Translation Start Codon (UCSC/hg19)

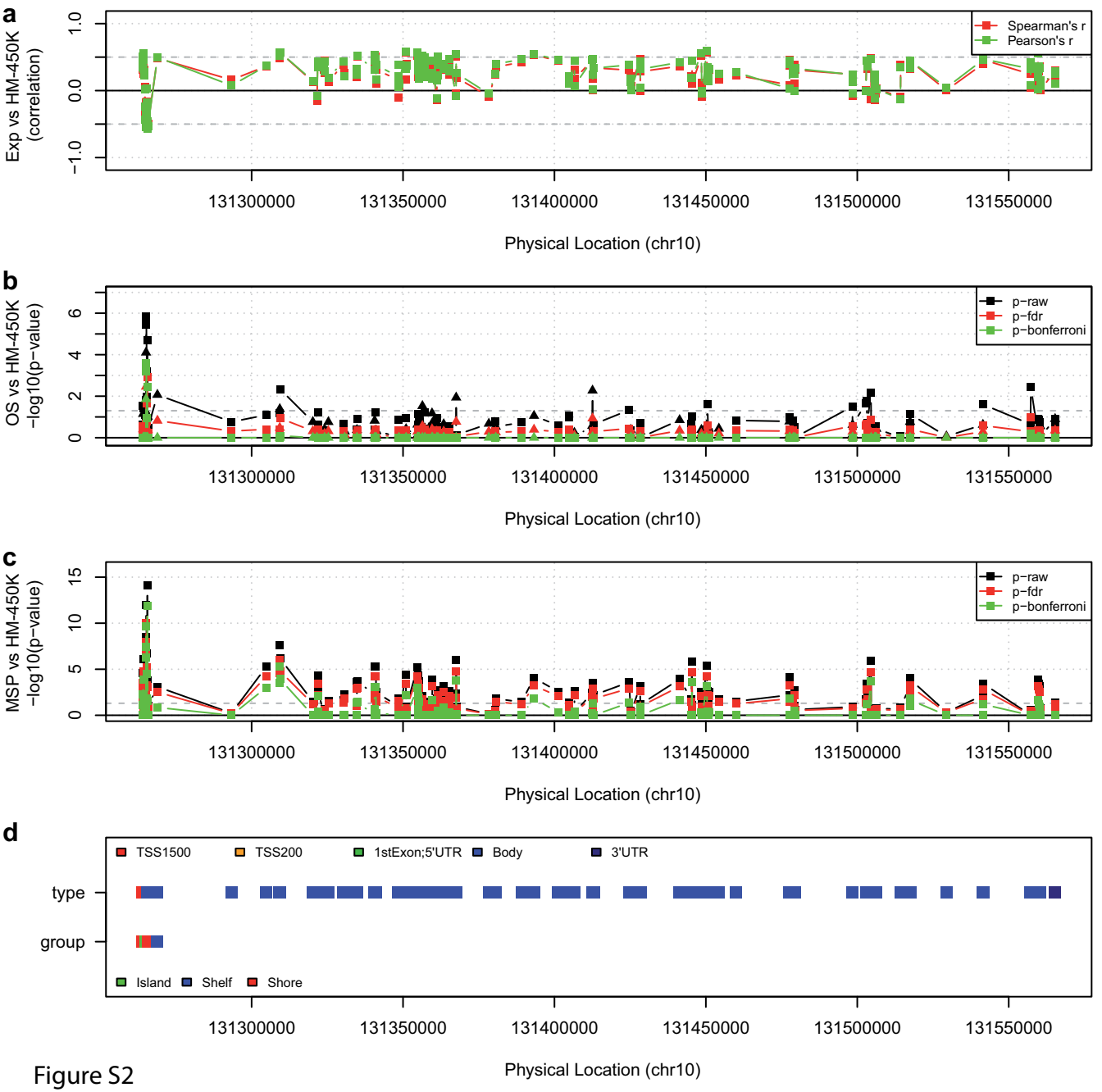

a
Exp vs HM-450K (correlation)
1.0
0.0
−1.0
Spearman's r
Pearson's r
131300000
131350000
131400000
131450000
131500000
131550000
Physical Location (chr10)
b
OS vs HM-450K −log10(p−value)
0
2
4
6
p−raw
p−fdr
p−bonferroni
131300000
131350000
131400000
131450000
131500000
131550000
Physical Location (chr10)
c
MSP vs HM-450K −log10(p−value)
0
5
10
15
p−raw
p−fdr
p−bonferroni
131300000
131350000
131400000
131450000
131500000
131550000
Physical Location (chr10)
d
TSS1500
TSS200
1stExon;5'UTR
Body
3'UTR
type
group
Island
Shelf
Shore
131300000
131350000
131400000
131450000
131500000
131550000
Physical Location (chr10)

Figure S2

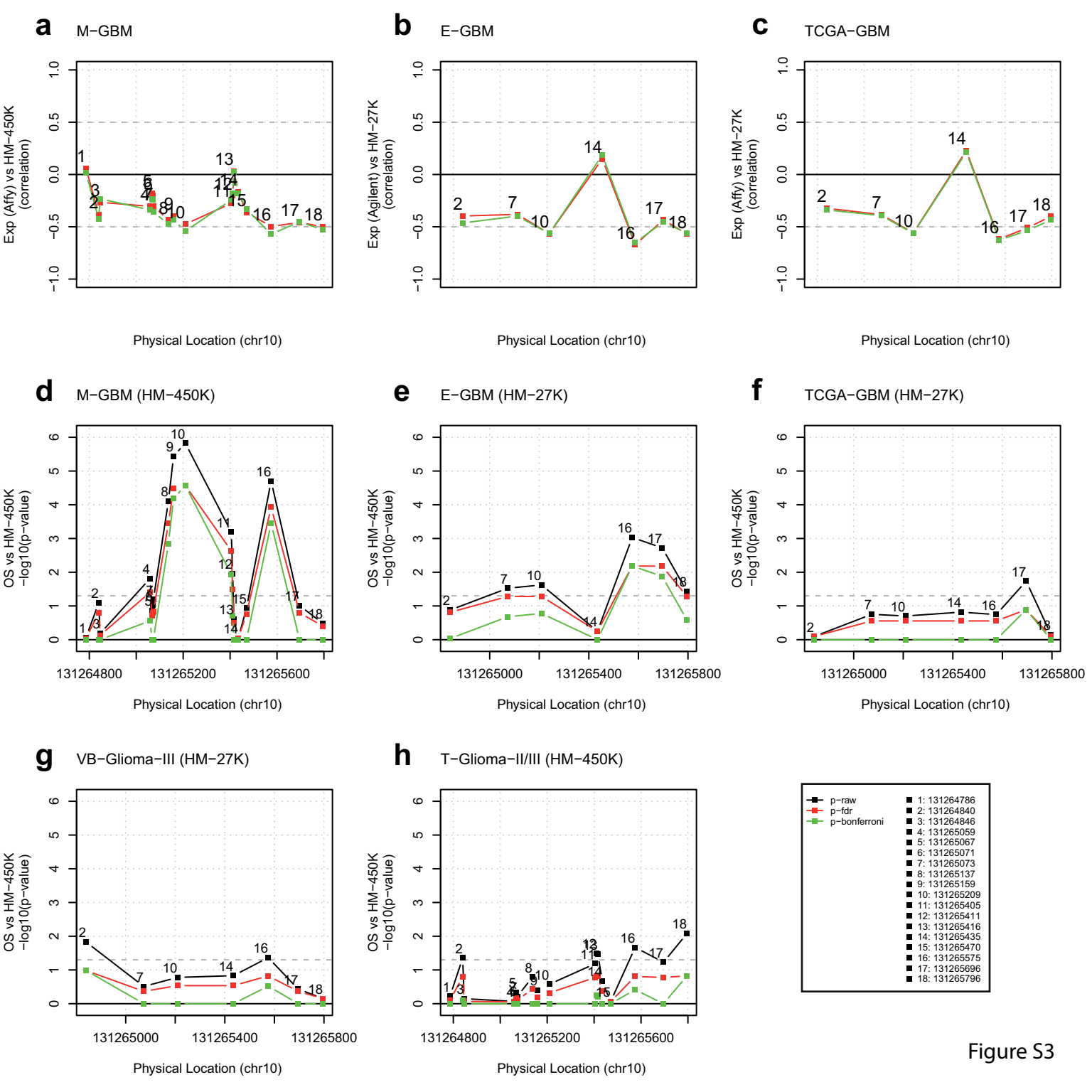

a
M−GBM
Exp (Affy) vs HM−450K (correlation)
Physical Location (chr10)
b
E−GBM
Exp (Agilent) vs HM−27K (correlation)
Physical Location (chr10)
c
TCGA−GBM
Exp (Affy) vs HM−27K (correlation)
Physical Location (chr10)
d
M−GBM (HM−450K)
OS vs HM−450K −log10(p−value)
131264800
131265200
131265600
Physical Location (chr10)
e
E−GBM (HM−27K)
OS vs HM−450K −log10(p−value)
131265000
131265400
131265800
Physical Location (chr10)
f
TCGA−GBM (HM−27K)
OS vs HM−450K −log10(p−value)
131265000
131265400
131265800
Physical Location (chr10)
g
VB−Glioma−III (HM−27K)
OS vs HM−450K −log10(p−value)
131265000
131265400
131265800
Physical Location (chr10)
h
T−Glioma−II/III (HM−450K)
OS vs HM−450K −log10(p−value)
131264800
131265200
131265600
Physical Location (chr10)
p−raw
p−fdr
p−bonferroni
1: 131264786
2: 131264840
3: 131264846
4: 131265059
5: 131265067
6: 131265071
7: 131265073
8: 131265137
9: 131265159
10: 131265209
11: 131265405
12: 131265411
13: 131265416
14: 131265435
15: 131265470
16: 131265575
17: 131265696
18: 131265796

Figure S3

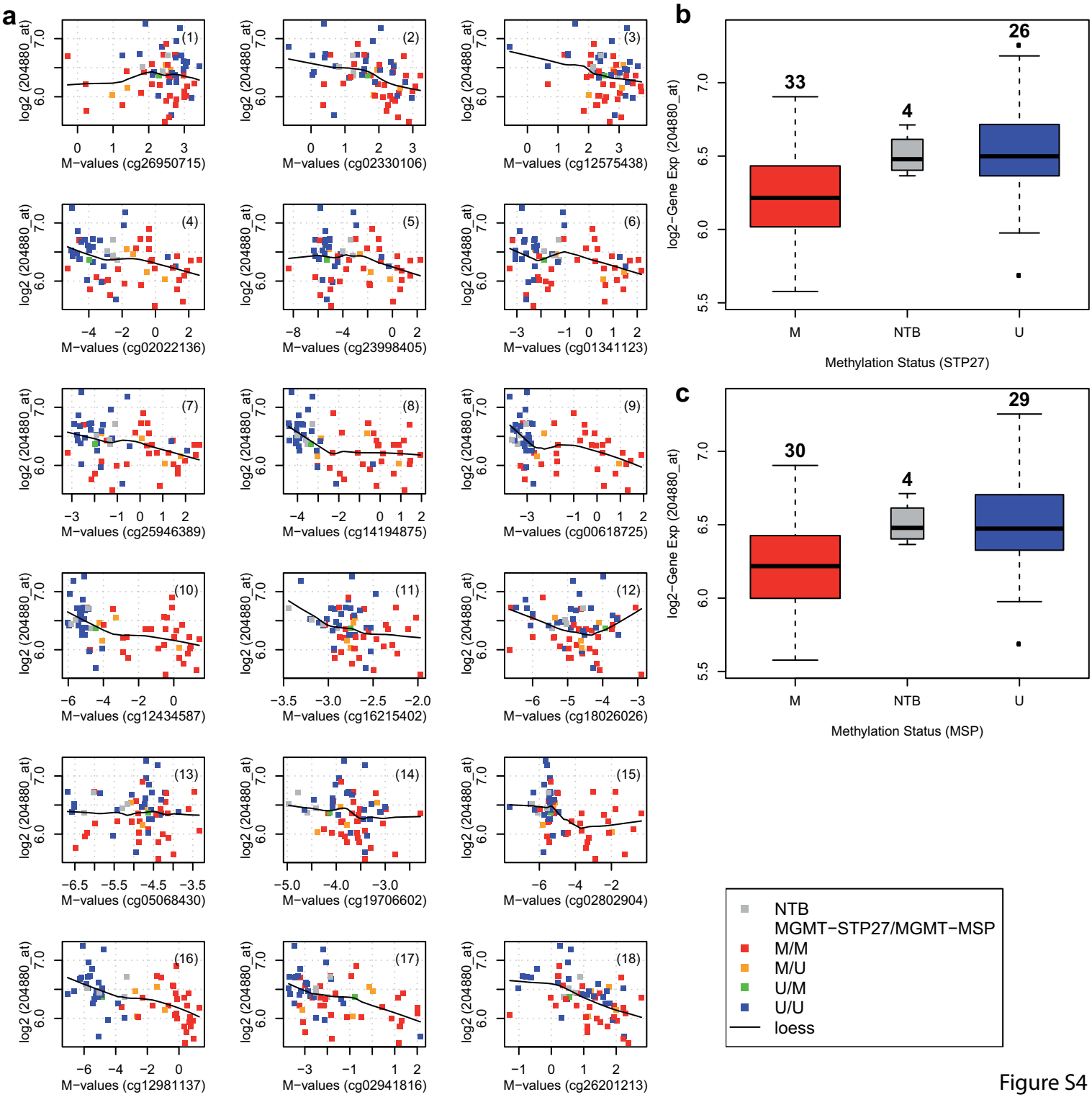
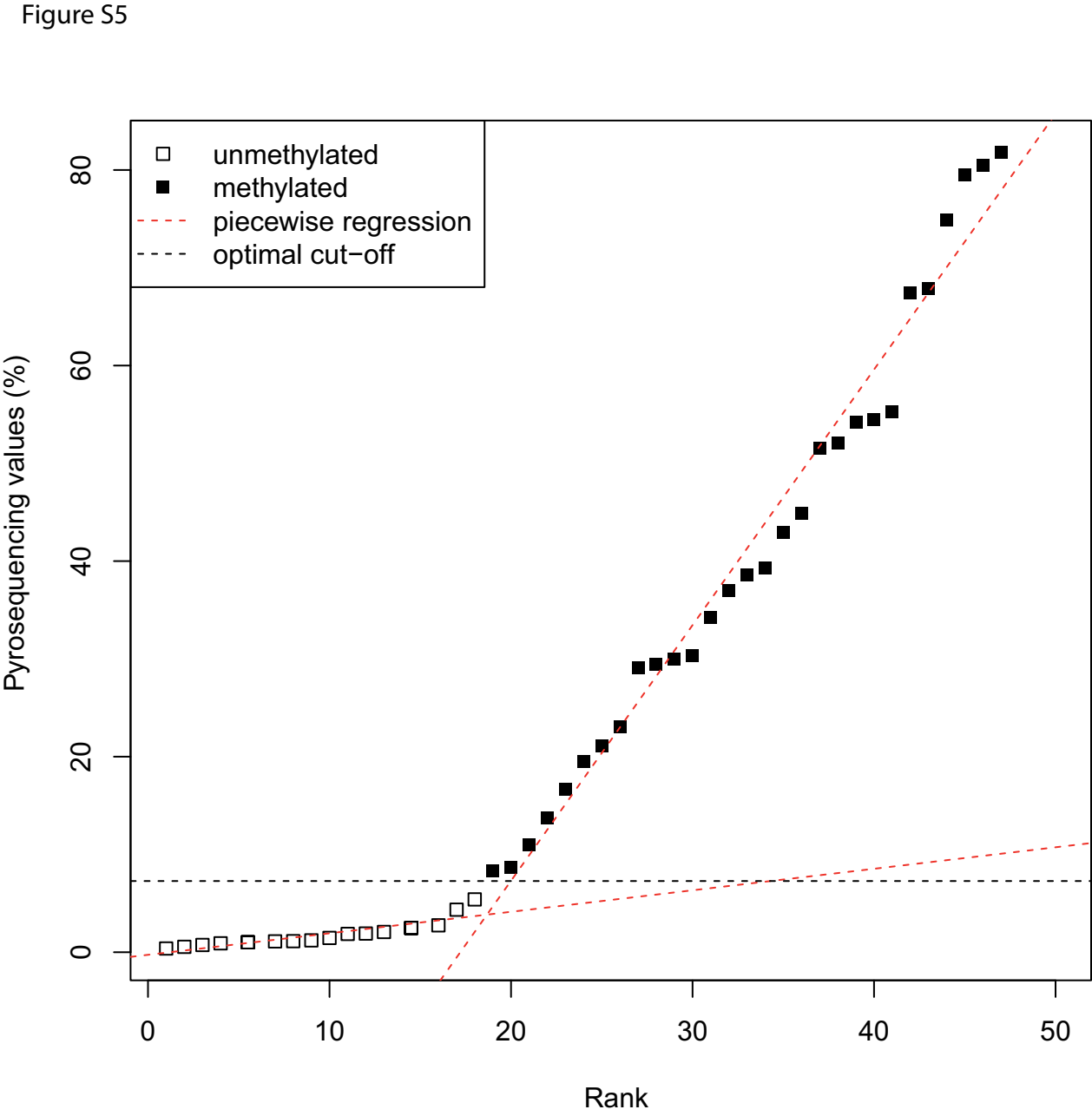

a
(1)
M−values (cg26950715)
(2)
M−values (cg02330106)
(3)
M−values (cg12575438)
(4)
M−values (cg02022136)
(5)
M−values (cg23998405)
(6)
M−values (cg01341123)
(7)
M−values (cg25946389)
(8)
M−values (cg14194875)
(9)
M−values (cg00618725)
(10)
M−values (cg12434587)
(11)
M−values (cg16215402)
(12)
M−values (cg18026026)
(13)
M−values (cg05068430)
(14)
M−values (cg19706602)
(15)
M−values (cg02802904)
(16)
M−values (cg12981137)
(17)
M−values (cg02941816)
(18)
M−values (cg26201213)
log2 (204880_at)
b
log2−Gene Exp (204880_at)
33
4
26
M
NTB
U
Methylation Status (STP27)
c
log2−Gene Exp (204880_at)
30
4
29
M
NTB
U
Methylation Status (MSP)
NTB
MGMT−STP27/MGMT−MSP
M/M
M/U
U/M
U/U
loess

Figure S4

Figure S5


unmethylated
methylated
piecewise regression
optimal cut−off
Pyrosequencing values (%)
0
20
40
60
80
0
10
20
30
40
50
Rank

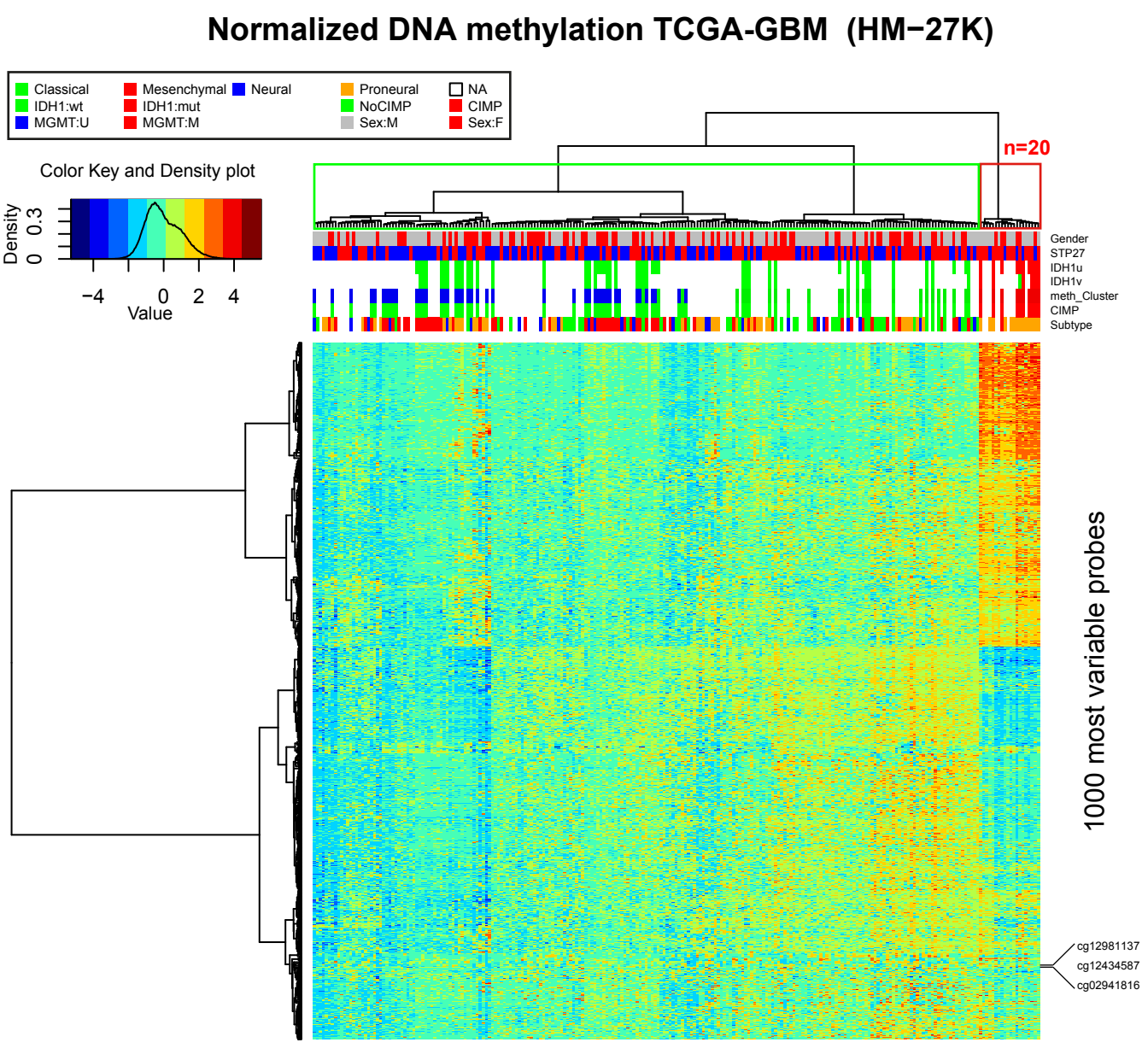
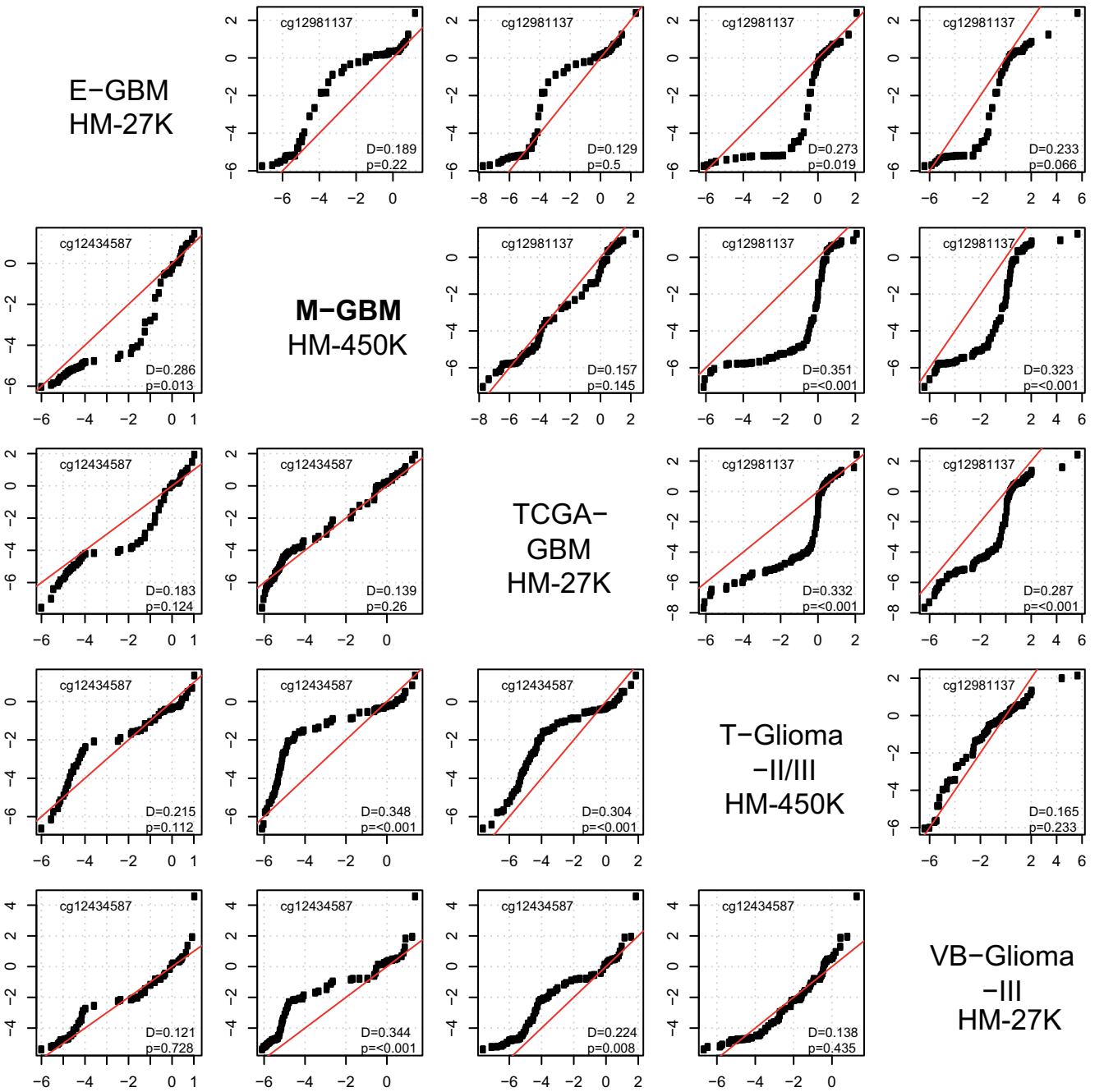

Normalized DNA methylation TCGA-GBM (HM−27K)
Classical
IDH1:wt
MGMT:U
Mesenchymal
IDH1:mut
MGMT:M
Neural
Proneural
NoCIMP
Sex:M
NA
CIMP
Sex:F
Color Key and Density plot
Density
0 0.3
−4 0 2 4
Value
n=20
Gender
STP27
IDH1u
IDH1v
meth_Cluster
CIMP
Subtype
1000 most variable probes
cg12981137
cg12434587
cg02941816
sample (n=241)

Figure S6

Figure S7


E−GBM
HM-27K
M−GBM
HM-450K
TCGA−
GBM
HM-27K
T−Glioma
−II/III
HM-450K
VB−Glioma
−III
HM-27K
cg12981137
cg12434587
D=0.189
p=0.22
D=0.129
p=0.5
D=0.273
p=0.019
D=0.233
p=0.066
D=0.286
p=0.013
D=0.157
p=0.145
D=0.351
p=<0.001
D=0.323
p=<0.001
D=0.183
p=0.124
D=0.139
p=0.26
D=0.332
p=<0.001
D=0.287
p=<0.001
D=0.215
p=0.112
D=0.348
p=<0.001
D=0.304
p=<0.001
D=0.165
p=0.233
D=0.121
p=0.728
D=0.344
p=<0.001
D=0.224
p=0.008
D=0.138
p=0.435